# مقدمة في الإحصاء للمهندسين الطبيين الحيويين

تأليف

كريستينا إم روبيلا

ترجمة

أ.د.م. حسن فؤاد محمد السيد

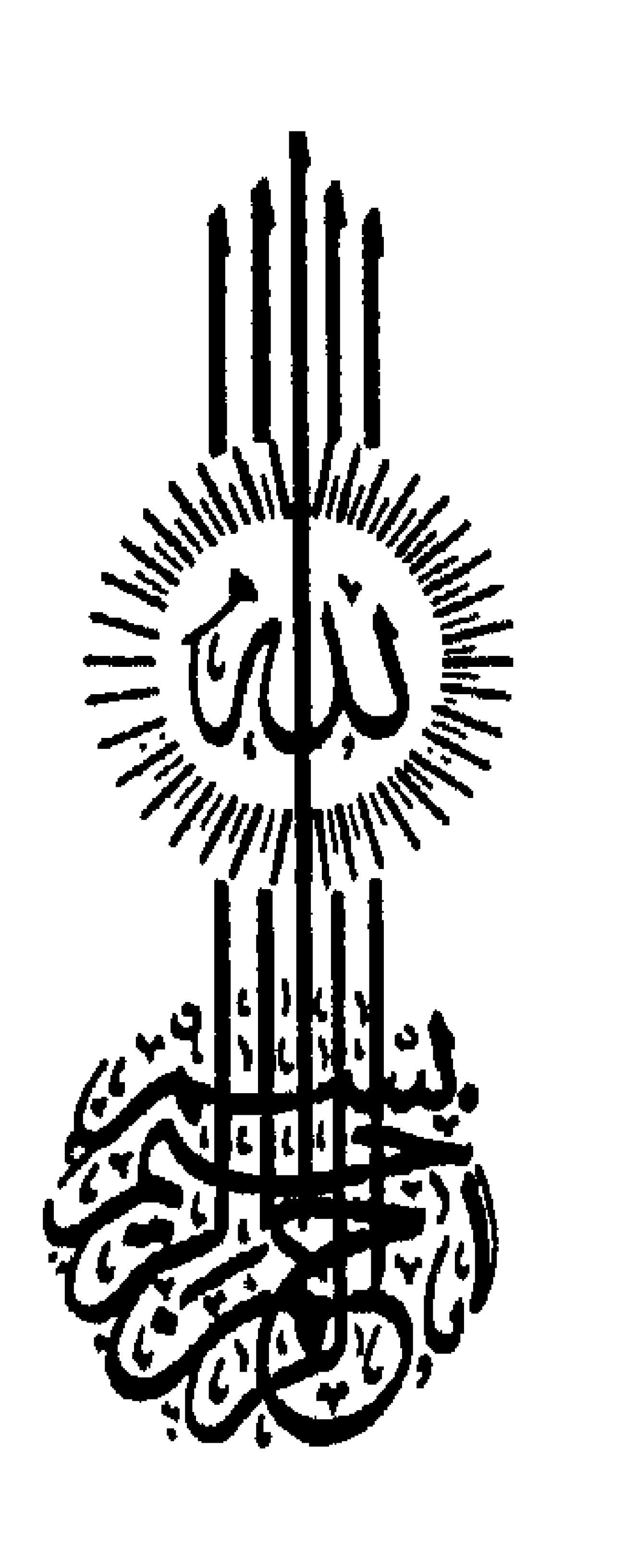
د. ممدوح منيف منيف





جامعة الملك سعود

النشر العلمي و المطابع



# مقدمة في الإحصاء للمهندسين الطبيين الحيويين


تأليف

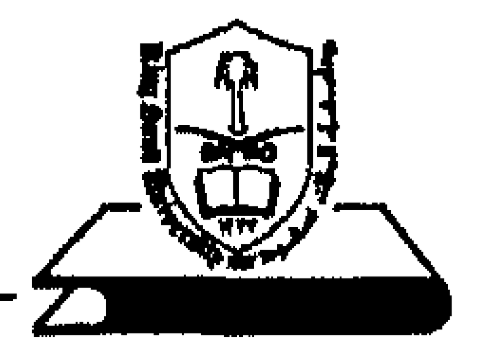
كريستينا إم روبيلا

ترجمة

أ. د. م. حسن فؤاد محمد السيد
أستاذ الهندسة الطبية

د. ممدوح منيف منيف
أستاذ الهندسة الطبية المساعد

كلية المجتمع – جامعة الملك سعود


النشر العلمي والمطابع - جامعة الملك سعود

ص.ب ٦٨٩٥٣ - الرياض ١١٥٣٧ - المملكة العربية السعودية

هذه ترجمة عربية مصرح بها من مركز الترجمة بالجامعة لكتاب
Introduction to Statistics for Biomedical Engineers
By: Kristina M. Ropella
© Morgan & Claypool, 2007

*فهرسة مكتبة الملك فهد الوطنية أثناء النشر*

إم روبيلا، كريستينا

مقدمة في الإحصاء للمهندسين الطبيين الحيويين./ كريستينا إم روبيلا؛ حسن فؤاد محمد السيد؛ ممدوح منيف منيف.- الرياض، ١٤٣٣هـ

١٣٦ص؛ ١٧×٢٤سم

ردمك: ٣-٠٠٠-٥٠٧-٦٠٣-٩٧٨

١- الإحصاء ٢- الطرق الإحصائية أ. السيد، حسن فؤاد محمد (مترجم)

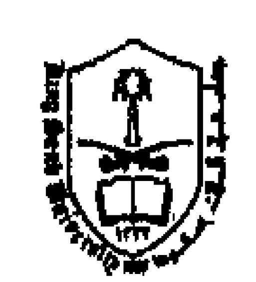
ب. منيف، ممدوح منيف (مترجم) ج. العنوان

ديوي ٤٢٢,٠٠١ ٣٤٨٥/ ١٤٣٣

رقم الإيداع: ٣٤٨٥/ ١٤٣٣

ردمك: ٣-٠٠٠-٥٠٧-٦٠٣-٩٧٨

حكمت هذا الكتاب لجنة متخصصة، وقد وافق المجلس العلمي على نشره في اجتماعه الحادي عشر للعام الدراسي ١٤٣٢/ ١٤٣٣هـ المعقود بتاريخ ٦/ ٣/ ١٤٣٣هـ الموافق ٢٩/ ١/ ٢٠١٢م.



النشر العلمي والمطابع ١٤٣٣هـ

# مقدمة المترجمين

مساهَمةً منا في إثراء المكتبة العربية بهذا النوع من الكتب المتخصصة والنوعية فقد قمنا بعون الله وتوفيقه بترجمة هذا الكتاب "مقدمة في الإحصاء للمهندسين الطبيين الحيويين" من أجل خدمة طلاب السنوات النهائية من مرحلة البكالوريوس والدراسات العليا في الهندسة الحيوية الطبية، والذي نرجو أن يستفيد منه الطلاب في دراستهم والمهندسون الطبيون في عملهم.

لطالما كانت كتب الإحصاء نظرية ومرعبة إلى حدٍّ ما لأولئك الذين لا يمارسون التحليل الإحصائي على أساس روتيني. وهكذا، فإن كثيراً من المهندسين والباحثين، الذين يحتاجون إلى استخدام الإحصاء بشكل متكرر، يفتقرون إلى المعرفة الكافية لاستخدام الإحصاء. يقدم هذا الكتاب المفردات والمفاهيم الأساسية للاحتمال والإحصاء المطلوبيْن لإجراء المراجعة العددية والتحليل الإحصائي المُستخدميْن في مجال الطب الحيوي. ويعتبر هذا الكتاب نقطة انطلاق لقضايا هامة يجب أخذها في الاعتبار عند تصميم التجارب، وتلخيص البيانات، وافتراض نموذج احتمال للبيانات، واختبار الفرضيات، واستخلاص النتائج من البيانات التي تم أخذها كعينات. من المفيد للقارئ الحصول على بعض المعلومات الأساسية في علوم الحياة وعلم وظائف الأعضاء وأن يكون على دراية بالأجهزة الطبية الحيوية الأساسية المُستخدمة في البيئة الإكلينيكية.

إن ترجمة الكتب العلمية إلى اللغة العربية مليئة بالمصاعب، حيث يجب الالتزام بالنص الأصلي على أن تكون الترجمة بلغة عربية سلسة وسليمة وتعكس المفهوم العلمي والمعنى الصحيح بصورة دقيقة ومفهومة. من أجل ذلك فقد قام المترجمان ببذل كل جهد ممكن لتحقيق جميع العناصر المذكور سابقًا ونرجو من الله أن نكون قد وفِّقنا في ذلك لخير وفائدة القراء والمهتمين بمواضيع هذا الكتاب.

وعرفانًا منا بالجميل نشكر جامعة الملك سعود ومركز الترجمة على تشجيع منسوبي الجامعة على التميز والريادة. ولا ننسى في هذا المجال أن نشكر المحكمين والمراجعين وكل من ساهم بشكل أو بآخر في إنجاز هذا العمل.

**المترجمان**

# إهداء

يتم تكريس هذا الكتاب لجميع طلابي الذين أنهوا مقرر الإحصاء BIEN 084 لمهندسي الطبية الحيوية وعلموني كيف أكون أكثر فعالية في نقل الموضوع وتقديم إحصائيات تتحقق بالنسبة لهم. وأشكر أيضاً (جي كليبول) J. Claypool على صبره وتشجيعه لي لإنهاء هذا الكتاب. وأخيراً، أشكر عائلتي لإتاحتهم الوقت لي في المنزل للعمل على الكمبيوتر المحمول.

# الملخص
# Abstract

هناك العديد من الكتب مكتوبة عن الإحصاء بعضها موجزة، وبعضها مفصَّلة، وبعضها ظريفة، وبعضها نابضة بالحيوية، وبعضها جافة تماماً. وقد تم تصميم كل كتاب من هذه الكتب لجمهور محدد. وفي كثير من الأحيان كانت كتب الإحصاء نظرية ومرعبة إلى حد ما لهؤلاء الذين لا يمارسون التحليل الإحصائي على أساس روتيني. وهكذا، فإن كثيراً من المهندسين والعلماء، الذين يحتاجون إلى استخدام الإحصاء بشكل متكرر أكثر من حساب التفاضل والتكامل أو المعادلات التفاضلية، يفتقرون إلى المعرفة الكافية لاستخدام الإحصاء. إن الجمهور الذي يتم تناوله في هذا الكتاب هو طالب الهندسة الحيوية الطبية على المستوى الجامعي والذي يحتاج إلى تغطية أساسية لمعظم التحليل الإحصائي الأساسي الذي يكثر استخدامه في ممارسة الهندسة الحيوية الطبية. يقدم الكتاب للطلاب المفردات والمفاهيم الأساسية للاحتمال والإحصاء المطلوبيْن لإجراء المراجعة العددية والتحليل الإحصائي المُستخدميْن في مجال الطب الحيوي. ويعتبر هذا الكتاب نقطة بداية لقضايا هامة يجب أخذها في الاعتبار عند تصميم التجارب، وتلخيص البيانات، وافتراض نموذج احتمال للبيانات، واختبار الفروض، واستخلاص النتائج من العينات المعطاة.

ينبغي للطالب الذي أكمل هذا الكتاب أن يكون لديه مفردات كافية لقراءة كتب أكثر تقدماً عن الإحصاء بالإضافة إلى معرفته بطرق التحليل العددي التي يتم استخدامها في مجال الهندسة الطبية الحيوية ولكنها خارج نطاق هذا الكتاب. تم تصميم هذا الكتاب لتكملة مقرر على مستوى المرحلة الجامعية في علم الإحصاء التطبيقي، وتحديداً في الهندسة الطبية الحيوية. وقد يستخدم المهندسون الممارسون الذين ليس لديهم دراسة رسمية في الإحصاء هذا الكتاب أيضاً باعتباره مقدمة بسيطة وقصيرة إلى الإحصاء المستخدم في الهندسة الطبية الحيوية. ينصب التركيز على تطبيق الإحصاء، والافتراضات الموضوعة في تطبيق الاختبارات الإحصائية، والقيود المفروضة على هذه الطرق الإحصائية الأولية، والأخطاء التي يتم ارتكابها في كثير من الأحيان باستخدام التحليل الإحصائي. ويتم تقديم عدد من الأمثلة مُستمدة من بحوث الهندسة الطبية الحيوية والممارسة الصناعية لمساعدة القارئ في فهم المفاهيم والتطبيق. من المفيد للقارئ الحصول على بعض المعلومات الأساسية في علوم الحياة وعلم وظائف الأعضاء (الفسيولوجي)، وأن يكون على دراية بالأجهزة الطبية الحيوية الأساسية المُستخدمة في البيئة الإكلينيكية.

**الكلمات الرئيسية**

نموذج احتمال، اختبار الفروض، الفسيولوجي، أنوفا ANOVA، التوزيع الطبيعي، فترات الثقة، اختبار القدرة

# المحتويات
# Contents

# الفصل الأول

## مقدمة
## Introduction

يجمع مهندسو الطبية الحيوية عادة جميع أنواع البيانات، من المرضى، والحيوانات، وعدادات الخلية، والمقايسات الصغرى، وأنظمة التصوير، ومبدِّلات الضغط، وأجهزة المراقبة بجانب السرير، وعمليات التصنيع، وأنظمة اختبار المواد، وأنظمة القياس الأخرى التي تدعم طيفاً واسعاً من البحوث والتصميم وبيئات التصنيع. في النهاية، فإن سبب تجميع البيانات هو اتخاذ قرار ما. وقد يهتم هذا القرار بتمييز الخصائص البيولوجية بين مختلف قطاعات السكان من الناس، أو تحديد ما إذا كان العلاج الدوائي فعالا، أو تحديد ما إذا كان فعالاً من حيث التكلفة الاستثمار بعدة ملايين من الدولارات في تكنولوجيا التصوير الطبي، أو تحديد ما إذا كانت عملية التصنيع تحت السيطرة، أو اختيار العلاج التأهيلي الأفضل لمريض معين.

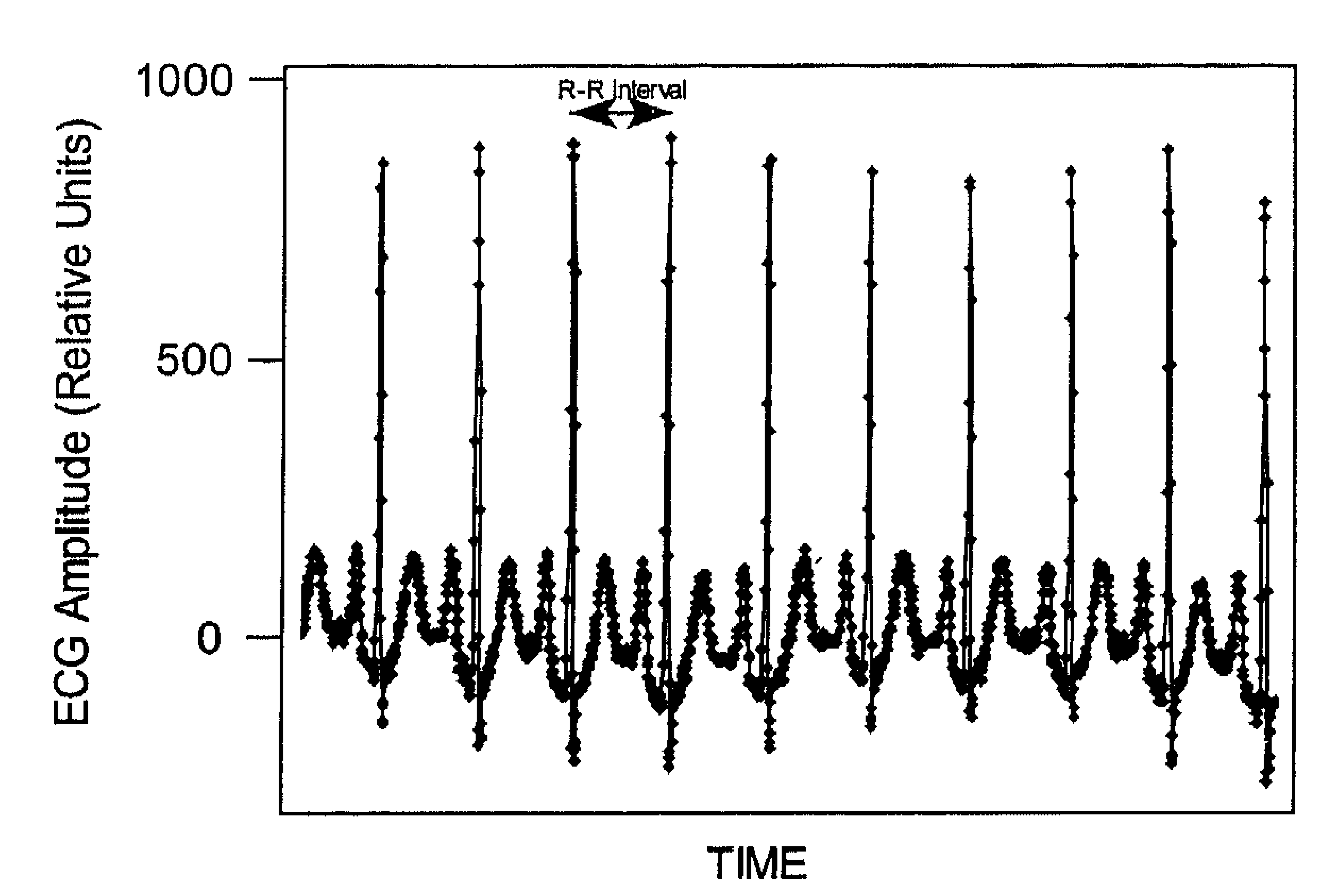
غالباً ما يكمن التحدي في صنع مثل هذه القرارات في حقيقة أن جميع البيانات في العالم الحقيقي تحتوي على بعض عناصر الشك بسبب العمليات العشوائية التي تكمن وراء معظم الظواهر الفيزيائية. تمنعنا هذه العناصر العشوائية من التنبؤ بالقيمة الدقيقة لأي كمية فيزيائية في أي لحظة من الزمن. وبعبارة أخرى، عندما نقوم بتجميع عينة أو نقطة بيانات، فإنه لا يمكننا عادة التنبؤ بالقيمة الدقيقة لتلك العينة أو النتيجة

التجريبية. على سبيل المثال، على الرغم من أن متوسط معدل ضربات القلب في حالة الراحة للبالغين العاديين حوالي 70 ضربة في الدقيقة، فإنه لا يمكننا التنبؤ بزمن الوصول الدقيق لضربة القلب التالية. ومع ذلك، يمكننا تقريب الاحتمال بأن زمن وصول الضربة التالية سيقع في فترة زمنية محددة إذا كان لدينا نموذج احتمال جيد لوصف الظاهرة العشوائية المساهمة في الفاصل الزمني بين ضربات القلب. يتأثر توقيت ضربات القلب بعدد من المتغيرات الفسيولوجية [1]، بما في ذلك فترة عدم الاستجابة الفسيولوجية للخلايا الفردية التي تشكّل عضلة القلب، وتسريب أغشية الخلايا في العقدة الجيبية (ناظم الخطى الطبيعي للقلب)، ونشاط الجهاز العصبي المستقل، الذي قد يُسرِّع أو يُبطئ معدل ضربات القلب كاستجابة لحاجة الجسم لزيادة تدفق الدم والأكسجين والمُغذِّيات. يؤدي جمع هذه العمليات البيولوجية إلى إنتاج شكل من ضربات القلب يمكننا قياسه عن طريق عد معدل النبض من شريان الرسغ أو الشريان السباتي أو من خلال البحث عن أشكال أمواج QRS محددة في المخطط الكهربائي للقلب [2]. على الرغم من أن هذا الجمع للأحداث يجعل من الصعب علينا التكهن بالضبط متى ستصل ضربة القلب الجديدة، فإنه يمكننا التخمين، بقدر لا ريب فيه من الثقة، متى ستصل الضربة التالية. وبعبارة أخرى، يمكننا أن نحدد احتمالاً من الاحتمالات بأن ضربة القلب التالية ستصل في فترة زمنية محددة. إذا كان علينا أخذ جميع أزمنة الوصول الممكنة في الاعتبار، وحددنا احتمالاً لأزمنة الوصول تلك، فسيكون لدينا نموذج احتمال لفترات ضربات القلب. وإذا استطعنا إيجاد نموذج احتمال لوصف احتمالات وقوع حدث معين أو نتيجة تجريبية، فإنه يمكننا استخدام الطرق الإحصائية لاتخاذ القرارات. تصف نماذج الاحتمال خصائص المجتمع الإحصائي أو الظاهرة التي يجري دراستها. بعد ذلك يستخدم التحليل الإحصائي هذه النماذج لمساعدتنا في اتخاذ القرارات بشأن المجتمعات الإحصائية أو العمليات.

إن الاستنتاجات التي قد يحصل عليها المرء باستخدام التحليل الإحصائي جيدة فقط بقدر جودة النموذج الأساسي الذي يتم استخدامه لوصف الظاهرة في العالم الحقيقي، مثل الفترة الزمنية الفاصلة بين ضربات القلب. على سبيل المثال، يُظهِر القلب الذي يعمل بشكل طبيعي تغيُّرًا كبيرًا في الفترات الزمنية الفاصلة من ضربة إلى ضربة (الشكل ١,١). يعكس هذا التغيّر جهد الجسم المستمر للحفاظ على التوازن بحيث يمكن للجسم الاستمرار في أداء وظائفه الأساسية، وإمداد الجسم بالأكسجين والمُغذّيات اللازمة ليعمل بشكل طبيعي. وقد ثبت من خلال البحوث الطبية الحيوية أن هناك ضياعًا في تغيّر معدل ضربات القلب مرتبطًا ببعض الأمراض مثل السكري ومرض نقص التروية القلبية (مرض القلب الإقفاري). يسعى الباحثون لتحديد ما إذا كان هذا الاختلاف في التغيّر كبيراً بين الأشخاص الطبيعيين والأشخاص الذين يعانون من مرض قلبي (أي، أنه بسبب بعض التغيير الكامن في البيولوجي وليس مجرد نتيجة للصدفة) وعما إذا كان يمكن استخدامه للتنبؤ بتطور المرض [1]. وسيلاحظ المرء أن نموذج الاحتمال يتغير نتيجة للتغيرات في الوظيفة أو العملية البيولوجية الكامنة. في حالة التصنيع فإن نموذج الاحتمال المُستخدم لوصف خرج عملية التصنيع قد يتغير كدالة لتشغيل الآلة أو التغيرات في بيئة التصنيع المحيطة، مثل درجة الحرارة، أو الرطوبة، أو الإنسان المُشغِّل.

تساعدنا الإحصائيات، بالإضافة إلى مساعدتنا في وصف نموذج الاحتمال المرتبط بظاهرة في العالم الحقيقي، على اتخاذ القرارات عن طريق منحنا الأدوات الكمية لاختبار الفروض. ونحن ندعو هذا "**بالإحصائيات الاستنتاجية**"، حيث تتيح لنا نتائج الاختبار الإحصائي التوصل إلى استنتاجات أو القيام باستدلالات حول واحدة أو أكثر من المجتمعات الإحصائية التي تم استجرار العينات منها. في معظم الأحيان،

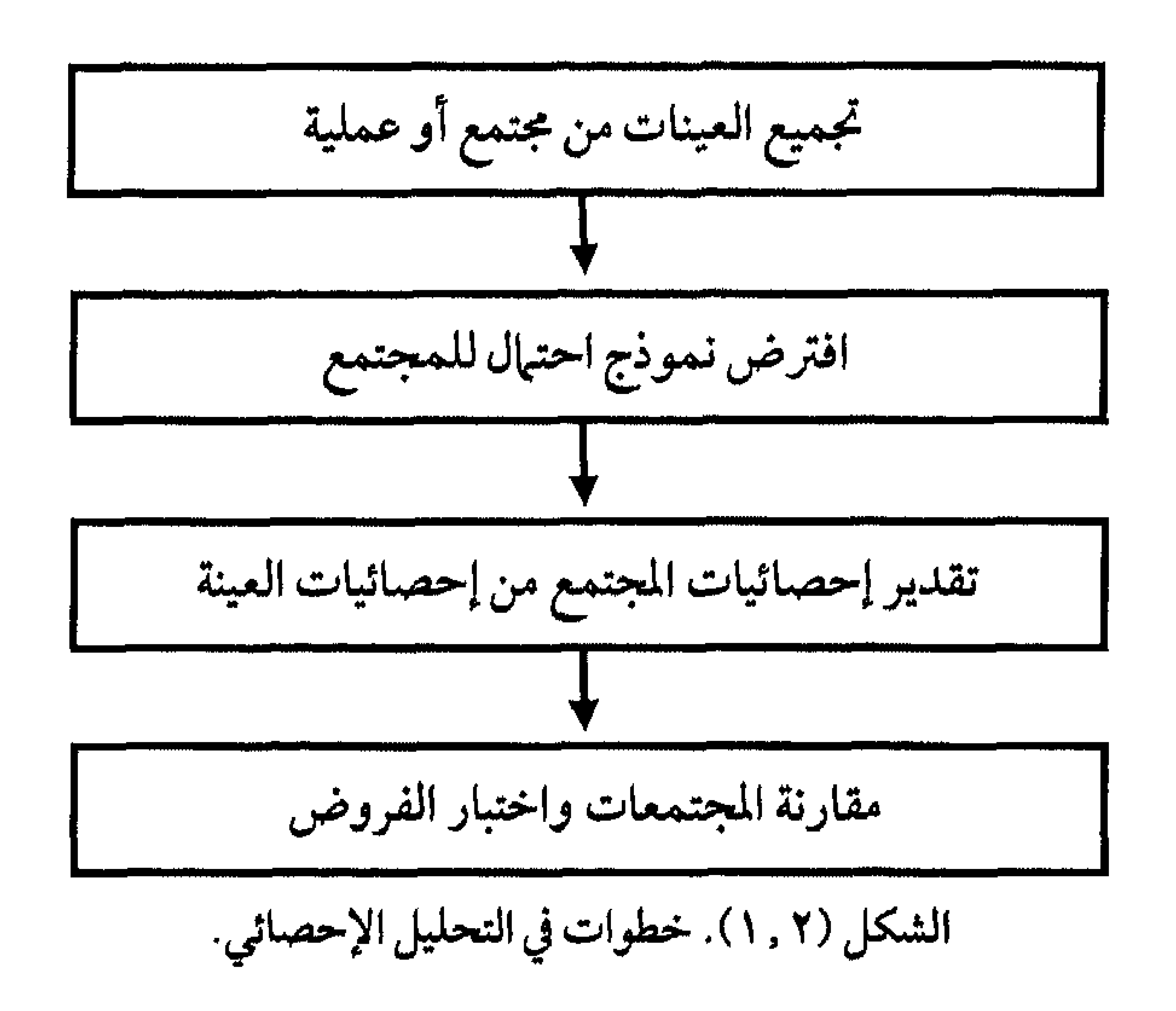
يهتم العلماء والمهندسون في مقارنة البيانات من مجتمعين أو أكثر من المجتمعات الإحصائية المختلفة أو من عمليتين أو أكثر من العمليات المختلفة. عادة ما تكون الفرضية الافتراضية هي أنه لا يوجد اختلاف في التوزيعات لمجتمعين أو أكثر من المجتمعات الإحصائية أو عمليتين أو أكثر من العمليات، ونحن نستخدم التحليل الإحصائي لتحديد ما إذا كانت هناك اختلافات حقيقية في توزيعات المجتمعات الإحصائية لضمان تحديد نماذج احتمال مختلفة للعمليات الفردية.




**الشكل (١ ، ١). مثال عن تسجيل الـ ECG (المخطط الكهربائي للقلب)، حيث يتم تحديد الفترة الفاصلة R-R كفترة زمنية فاصلة بين أمواج R المتتالية للمركبة QRS، وهو شكل الموجة الأكثر وضوحاً للـ ECG.**

وباختصار، عادة ما يجمع مهندسو الطبية الحيوية البيانات أو العينات من ظواهر مختلفة تحتوي على بعض عناصر التغيرات العشوائية أو العناصر التي لا يمكن التنبؤ بها، لأغراض اتخاذ القرارات. ولاتخاذ القرارات السليمة في سياق عدم اليقين مع مستوى

معين من الثقة ، فإننا بحاجة لافتراض بعض نماذج الاحتمال للمجتمعات الإحصائية التي تم جمع العينات منها. وحالما يتم افتراض النموذج الأساسي ، فإنه يمكننا تحديد الاختبارات الإحصائية المناسبة لمقارنة مجتمعين إحصائيين أو أكثر ومن ثم استخدام هذه التجارب لاستخلاص استنتاجات حول الفرضيات التي جمعنا البيانات لأجلها في المقام الأول. يحدد الشكل (١,٢) خطوات إجراء التحليل الإحصائي للبيانات.




الشكل (٢ , ١). خطوات في التحليل الإحصائي.

سنقوم في الفصول التالية بشرح طرق لتلخيص البيانات التي تم جمعها بيانياً وعددياً. ثم سنتحدث عن توافق نموذج احتمال للبيانات التي تم جمعها من خلال وصف موجز لعدد من نماذج الاحتمال المعروفة جيداً التي يتم استخدامها لوصف ظاهرة بيولوجية. وأخيراً، حالما يتم افتراض نموذج للمجتمعات التي جمعنا منها بيانات العينات التي لدينا ، فإننا سوف نناقش أنواع الاختبارات الإحصائية التي يمكن استخدامها لمقارنة البيانات من مجموعات متعددة وتسمح لنا باختبار الفروض حول المجتمعات الإحصائية الكامنة.

# الفصل الثاني

## تجميع البيانات والتصميم التجريبي
## Collecting Data and Experimental Design

قبل مناقشة أي نوع من ملخص البيانات والتحليل الإحصائي لها، من المهم أن ندرك أن قيمة أي تحليل إحصائي جيدة فقط بمقدار جودة البيانات التي تم جمعها. وبما أننا نستخدم بيانات أو عينات لاستخلاص استنتاجات حول مجتمعات أو عمليات كاملة، فمن الأهمية بمكان أن تكون البيانات التي تم جمعها (أو العينات التي تم جمعها) ممثلة للمجتمع الإحصائي الأكبر. وبعبارة أخرى، إذا كنا نحاول تحديد ما إذا كان الرجال الذين تتراوح أعمارهم بين 20 و50 سنة يستجيبون بشكل إيجابي لدواء يخفّض مستوى الكولسترول، فإننا بحاجة إلى اختيار دقيق لمجتمع الأشخاص التي نعطيها الدواء ونأخذ القياسات منها. بعبارة أخرى، يجب أن يكون لدينا ما يكفي من العينات لتمثيل التباينات ضمن المجتمع الإحصائي. هناك قدر كبير من الاختلاف في الوزن، والطول، والتركيب الجيني، والنظام الغذائي، وعادات ممارسة الرياضة، وتعاطي الدواء في جميع أعمار الرجال من 20 — 50 سنة والذين قد يكون لديهم أيضاً نسبة كولسترول عالية. إذا أردنا اختبار فعالية العقار الجديد في خفض نسبة الكولسترول، فإنه يجب علينا جمع ما يكفي من البيانات أو العينات للحصول على اختلاف التركيب البيولوجي والبيئة البيولوجية للمجتمع التي نهتم في علاجها بالدواء الجديد. إن

الحصول على هذا الاختلاف غالباً ما يكون التحدي الأكبر الذي يواجه مهندسي الطبية الحيوية في جمع البيانات واستخدام الإحصائيات لاستخلاص استنتاجات ذات مغزى. ويجب أن يطرح القائم بالتجربة أسئلة مثل ما يلي :

- ما نوع الشخص، الشيء، أو الظاهرة التي آخذ عيناتها؟
- ما المتغيرات التي أستطيع التحكم بها والتي تؤثر على القياس أو البيانات؟
- ما عدد العينات التي أحتاج إليها للحصول على متغيّر المجتمع الإحصائي لتطبيق الإحصائيات المناسبة واستخلاص استنتاجات ذات مغزى؟
- كيف يمكنني تجنب انحراف البيانات من خلال التصميم؟

إن التصميم التجريبي، على الرغم من أنه ليس الركيزة الأساسية لهذا الكتاب، هو الخطوة الأكثر أهمية لدعم التحليل الإحصائي الذي من شأنه أن يؤدي إلى استنتاجات ذات معنى، وبالتالي اتخاذ قرارات سليمة.

إن أحد أكثر الأسئلة الأساسية التي يطرحها باحثو الطبية الحيوية هو: "ما حجم العينة التي أحتاجها؟" أو "ما هو عدد الأشخاص الذين أحتاج إليهم لاتخاذ القرارات بأي مستوى من الثقة؟" سوف نعالج هذه الأسئلة المهمة في نهاية هذا الكتاب بعد أن يكون قد تم بالفعل تغطية مفاهيم مثل التباين، ونماذج الاحتمالات، واختبار الفروض. على سبيل المثال، سيتم وصف اختبارات القدرة كمتوسطات للتنبؤ بحجم العينة المطلوبة للكشف عن اختلافات كبيرة في متوسطات مجتمعين باستخدام الاختبار t.

هناك عنصران من عناصر التصميم التجريبي يعتبران حاسمين لمنع انحراف البيانات أو اختيار العينات التي لا تمثل إلى حد ما المجتمع الإحصائي، هما العشوائية (randomization) والحجب (blocking).

يشير التوزيع العشوائي إلى العملية التي نختار بواسطتها العينات أو الوحدات التجريبية عشوائياً من المجتمع الإحصائي الأكبر بحيث نزيد إلى الحد الأقصى من فرصتنا

في الحصول على التباين في المجتمع الإحصائي. وبعبارة أخرى، فإننا لا نحد العينات التي لدينا بحيث لا يتم الحصول في العينات إلا على جزء من خصائص أو سلوكيات المجتمع الإحصائي. والأهم من ذلك، إننا لا نحرِّف النتائج من خلال الحد بشكل مصطنع من التباين في العينات بحيث نعمل على تغيير نموذج الاحتمال لمجتمع العينات فيما يتعلق بنموذج الاحتمال للمجتمع الإحصائي.

بالإضافة إلى جعل اختيارنا للوحدات التجريبية، التي يتم من خلالها أخذ العينات، عشوائياً، فقد نجعل أيضاً واجب المعالجات لدينا عشوائياً بالنسبة للوحدات التجريبية لدينا. أو، قد نجعل الترتيب، الذي نأخذ البيانات به من الوحدات التجريبية، عشوائياً. على سبيل المثال، إذا كنا نختبر الفعالية لطريقتين مختلفتين من طرق التصوير الطبي لاكتشاف ورم في الدماغ، فسوف نحدد عشوائياً جميع الأشخاص الذين يشتبه بوجود ورم في الدماغ لديهم إلى إحدى طريقتي التصوير. وهكذا، إذا كان لدينا مزيج من الجنس، والعمر، والنوع من ورم الدماغ المشارك في الدراسة، فإننا نخفِّض من فرصة وجود مجتمع جميعه من جنس واحد أو مجتمع جميعه ذو عمر واحد مخصص لإحدى طرق التصوير ونوع مختلف جدًّا لمجتمع مخصص لطريقة التصوير الثانية. إذا تم ملاحظة فرق في نتائج طريقتي التصوير، فإننا لا نعرض الجنس أو العمر بشكل مصطنع كأحد العوامل التي تؤثر على نتائج التصوير.

وكمثال آخر، إذا كان المرء يختبر القوة لثلاث مواد مختلفة لاستخدامها في زراعة الورك باستخدام عدة مقاييس قوة من جهاز اختبار المواد، فقد يجعل المرء الترتيب، الذي يتم به تقديم العينات لمواد الاختبار الثلاثة المختلفة إلى الجهاز، عشوائياً. يمكن أن يتغير أداء الجهاز مع مرور الوقت بسبب الاستهلاك، ودرجة الحرارة، والرطوبة، والتشوه، والإجهاد، وخصائص المستخدم. إذا تم الطلب من المهندس الطبي الحيوي إيجاد المادة الأقوى من أجل ورك اصطناعي باستخدام معايير قوة محددة، فقد يقوم

بإجراء تجربة. دعونا نفترض أنه تم إعطاء المهندس ثلاثة صناديق، وكل صندوق يحتوي على خمس زرعات (implants) اصطناعية للورك مصنوعة من إحدى المواد الثلاثة: التيتانيوم، والفولاذ، والبلاستيك. وبالنسبة لأي صندوق كان، فإن عينات الزرع الخمسة جميعها مصنوعة من نفس المادة. ولاختبار قوة المادة للـ 15 زرعة المختلفة، فقد يجعل المهندس الترتيب، الذي يتم فيه اختبار كل واحدة من الزرعات الـ 15 في جهاز اختبار المواد، عشوائياً بحيث أن تغيرات أداء الجهاز التي تعتمد على الوقت أو التفاعلات بين الجهاز والمادة أو الظروف البيئية المتغيرة مع الزمن لا تؤدي إلى انحراف النتائج لواحدة أو أكثر من المواد. وهكذا، لجعل اختبار الزرعة عشوائياً تماماً، فقد يقوم المهندس حرفياً بوضع الأرقام من 1– 15 في قبعة وأيضاً تخصيص الأرقام من 1– 15 لكل زرعة يجب اختبارها. يقوم المهندس بعد ذلك بسحب أحد الأرقام من الـ 15 رقمًا لا على التعيين من القبعة واختبار الزرعة التي تتوافق مع هذا الرقم. بهذه الطريقة لا يختبر المهندس جميع المواد ذات النوع الواحد في أي ترتيب معين، وهكذا نتجنب إدخال آثار الترتيب إلى البيانات.

أما الجانب الثاني من التصميم التجريبي فهو الحجب. نهتم في العديد من التجارب بعامل أو عاملين أو متغير أو متغيرين محددين قد يؤثران على قياسنا أو عينتنا. ومع ذلك، قد يكون هناك عوامل أخرى تؤثر أيضاً على قياسنا وتربك إحصائياتنا. نحاول في التصميم التجريبي الجيد جمع العينات بحيث لا يتم حرف المعالجات المختلفة داخل العامل ذي الأهمية من خلال اختلاف قيم العوامل المربكة. وبعبارة أخرى، ينبغي أن نكون على يقين من أن كل معالجة ضمن العامل ذي الأهمية لدينا يتم اختبارها ضمن كل قيمة لعامل الإرباك. ونشير إلى هذا التصميم على أنه الحجب بواسطة عامل الإرباك. على سبيل المثال، قد نرغب في دراسة فقدان الوزن كتابع لثلاث حبوب حمية مختلفة. إن أحد العوامل المربكة قد يكون بداية وزن الشخص. وهكذا، قد نريد في اختبار فعالية الحبوب الثلاثة في خفض الوزن، حجب الأشخاص

وفقاً لبداية الوزن. وهكذا، قد نقوم أولاً بتصنيف الأشخاص وفقاً لبداية أوزانهم ومن ثم اختبار كل حبة من حبوب الحمية ضمن كل مجتمع لبداية الأوزان.

نقوم في البحوث الطبية الحيوية بالحجب بواسطة الوحدة التجريبية في كثير من الأحيان. عندما يكون هذا النوع من الحجب جزءًا من التصميم التجريبي، فإن القائم بالتجربة يجمع عينات متعددة من البيانات، حيث تمثل كل عينة ظروفًا تجريبية مختلفة، من كل وحدة من الوحدات التجريبية. يوفّر الشكل (٢،١) مخططاً لتجربة يتم فيها جمع البيانات قبل وبعد تلقي المرضى للعلاج، ويستخدم التصميم التجريبي الحجب (اليسار) أو عدم الحجب (اليمين) بواسطة وحدة تجريبية. في حالة الحجب، يتم جمع البيانات قبل وبعد العلاج من نفس مجتمع الأشخاص. وهكذا، داخل الفرد، فإن نفس العوامل البيولوجية التي تؤثر على الاستجابة البيولوجية للعلاج موجودة قبل وبعد العلاج. إن كل شخص هو بمثابة تحكم بذاته للعوامل التي قد تتغير بشكل عشوائي من شخص إلى شخص قبل وبعد العلاج على حد سواء. نتخلص في الجوهر من الانحرافات في الاختلافات التي قد تنتج بين المجموعتين (قبل وبعد)؛ لأننا نستخدم مجموعتين مختلفتين من الوحدات التجريبية. على سبيل المثال، إذا استخدمنا مجتمعًا واحدًا من الأشخاص قبل العلاج ومن ثم مجتمعًا مختلفًا تماماً من الأشخاص بعد العلاج (الشكل ٢،١، اليمين)، فإن هناك فرصة بأن مجموعتي الأشخاص قد تختلفان بما فيه الكفاية في الجنس، أو العمر، أو الوزن، أو العرق، أو التركيب الجيني، مما قد يؤدي إلى اختلاف في الاستجابة للعلاج الذي له علاقة بسيطة بالعلاج الأساسي. وبعبارة أخرى، قد تكون هناك عوامل إرباك تساهم في اختلاف النتيجة التجريبية قبل وبعد العلاج والتي ليست فقط عاملا للعلاج ولكنها في الحقيقة تشويش صنعي للاختلافات في توزيع المجموعتين المختلفتين للأشخاص الذين تم اختيار مجموعتي العينات منهم. يساعد الحجب على التخلص من تأثير التغير داخل الشخص (intersubject).

| حجب (قياسات مُكرَّرة) | | | | عدم حجب (قياسات غير مُكرَّرة) | | | |
|---|---|---|---|---|---|---|---|
| الموضوع | القياس قبل المعالجة | القياس بعد المعالجة | | الموضوع | القياس قبل المعالجة | الموضوع | القياس بعد المعالجة |
| 1 | M11 | I2 | | 1 | M1 | K+1 | M(K+1) |
| 2 | M21 | M22 | | 2 | M2 | K+2 | M(K+2) |
| 3 | M31 | M32 | | 3 | M3 | K+3 | M(K+3) |
| . | | | | . | | . | |
| . | | | | . | | . | |
| K | MK1 | MK2 | | K | MK | K+K | M(K+K) |

**الشكل (١ ,٢). عينات تم استخلاصها من مجتمعين (قبل وبعد العلاج)، ويستخدم التصميم التجريبي الحجب (اليسار) أو عدم الحجب (اليمين). في هذه الحالة فإن الحجب هو الوحدة التجريبية (الشخص) التي تم منها إجراء القياس.**

ومع ذلك، فإن الحجب ليس دائماً ممكناً، نظراً لطبيعة بعض دراسات البحوث الطبية الحيوية. على سبيل المثال، إذا أراد المرء دراسة فعالية دوائين مختلفين للعلاج الكيميائي للتقليل من حجم الورم، فإنه من غير العملي اختبار الدوائين معاً على كتلة الورم نفسه. وهكذا، يتم اختبار الدوائين على مجموعتين مختلفتين من الأفراد. وقد يكون نفس نوع التصميم ضرورياً لاختبار فعالية أنظمة فقدان الوزن

وهكذا، فإن بعض المفاهيم والتعاريف المهمة التي يجب أخذها في الاعتبار عند تصميم التجارب تشمل ما يلي:

- **الوحدة التجريبية:** المادة، أو الشيء، أو الشخص الذي نطبق العلاج عليه والذي نأخذ منه قياسات العينات.
- **العشوائية:** تخصيص العلاجات عشوائياً للوحدات التجريبية.
- **الحجب:** تعيين كل العلاجات داخل العامل لكل مستوى من مستويات عامل الحجب. وغالباً ما يكون عامل الحجب هو الوحدة التجريبية. لاحظ

أنه عند استخدام الحجب، فإننا لا نزال نجعل الترتيب الذي يتم به تطبيق العلاج لكل وحدة تجريبية عشوائياً لتجنب ترتيب الحجب.

وأخيراً، يجب على القائم بالتجربة أن يفكر دائماً حول كيفية تمثيل مجتمع العينة الكامنة بالنسبة للمجتمع الأكبر. بما أنه من المستحيل عملياً اختبار كل فرد من المجتمع الإحصائي أو كل منتج موجود على خط التجميع، وخصوصاً عندما يتم استخدام طرق الاختبار المدمرة، فإنه يجب على المهندس الطبي الحيوي في كثير من الأحيان جمع البيانات من عينة صغيرة جداً تم استخلاصها من المجتمع الإحصائي الأكبر. ومن المهم، إذا كانت الإحصائيات سوف تؤدي إلى استنتاجات مفيدة، أن يلتقط مجتمع العينات تباين المجتمع الإحصائي الكامن. ما هو حتى أكثر تحدياً هو أننا في كثير من الأحيان ليس لدينا فهماً جيداً لتباين المجتمع الإحصائي الكامن، وبسبب النفقات واحترام الحياة، فإننا عادة مقيَّدين بعدد العينات التي قد نجمعها من البحوث الطبية الحيوية والتصنيع. وهذه القيود ليست سهلة المعالجة وتحتاج إلى أن يأخذ المهندس دائماً في الاعتبار مقدار ملاءمة العينة وتحليل البيانات ومقدار جودة تمثيل المجتمع (المجتمعات) الإحصائية التي من خلالها يتم استخلاص العينات.

# الفصل الثالث

## ملخص البيانات والإحصاء الوصفي

## Data Summary and Descriptive Statistics

نفترض الآن أننا جمعنا بياناتنا من خلال استخدام تصميم تجريبي جيد. لدينا الآن مجتمع إحصائي من الأرقام، أو الملاحظات، أو توصيفات لوصف بياناتنا، ونود تلخيص البيانات لاتخاذ قرارات، أو اختبار فرضية، أو استخلاص نتيجة.

### (٣,١) لماذا نجمع البيانات؟

### WHY DO WE COLLECT DATA?

العالم مليء بالشك، بمعنى أن هناك عوامل عشوائية أو لا يمكن التنبؤ بها تؤثر على كل قياس تجريبي نقوم به. تنشأ الجوانب غير المتوقعة للنتائج التجريبية أيضاً من المتغيّر في الأنظمة البيولوجية (نتيجة لعوامل وراثية وبيئية) وعمليات التصنيع، والخطأ البشري في إجراء القياسات، والعمليات الأساسية الأخرى التي تؤثر على القياسات التي يتم إجراؤها.

على الرغم من عدم اليقين بشأن النتيجة الدقيقة لتجربة أو حدوث لحادثة في المستقبل، فإننا نجمع البيانات في محاولة لفهم أفضل للعمليات أو المجتمعات الإحصائية التي تؤثر على النتيجة التجريبية حتى نتمكن من القيام ببعض التوقعات. توفّر البيانات معلومات

لتقليل عدم اليقين، والسماح باتخاذ القرارات. وأيضا تجميع وتحليل البيانات يساعدنا في حل المشاكل بشكل صحيح. ولا يمكن التأكيد بما فيه الكفاية على أنه يجب تجميع البيانات وتحليلها بشكل صحيح إذا كان لتحليل البيانات والاستنتاجات اللاحقة أية قيمة.

## (٣,٢) لماذا نحن بحاجة إلى الإحصاء؟
## WHY DO WE NEED STATISTICS?

لدينا ثلاثة أسباب رئيسية لاستخدام وتحليل البيانات الإحصائية:

١- العالم الحقيقي مليء بالأحداث العشوائية التي لا يمكن وصفها بالصيغ الرياضية الدقيقة.

٢- المتغير هو صفة طبيعية وعادية للعالم الطبيعي.

٣- نود اتخاذ قرارات تتسم ببعض الثقة. وهذا يعني أننا في حاجة إلى إيجاد ميول ضمن المتغير.

## (٣,٣) ما هي المسائل التي نأمل معالجتها بالتحليل الإحصائي؟
## WHAT QUESTIONS DO WE HOPE TO ADDRESS WITH OUR STATISTICAL ANALYSIS?

هناك العديد من المسائل الأساسية نأمل في معالجتها عند استخدام موجز عددي وبياني للبيانات:

١- هل نستطيع المفاضلة بين المجتمعات الإحصائية أو الجماعات؟

٢- هل هناك ارتباطات بين المتغيرات أو المجتمعات الإحصائية؟

٣- هل العمليات تحت السيطرة؟

إن إيجاد الاختلافات الفسيولوجية بين المجتمعات الإحصائية هو على الأرجح الهدف الأكثر شيوعاً للبحوث الطبية الحيوية. على سبيل المثال، قد يريد الباحثون

معرفة إذا كان هناك اختلاف في متوسط العمر المتوقع بين الناس الذين يعانون من زيادة الوزن والناس الذين يعانون من نقص الوزن. أو قد تريد شركة دوائية تحديد ما إذا كان أحد أنواع المضادات الحيوية أكثر فعالية في مكافحة البكتيريا من نوع آخر. أو يتساءل طبيب عما إذا كان يتم تخفيض ضغط الدم الانبساطي في مجتمع إحصائي من أشخاص يعانون ارتفاع ضغط الدم بعد تناول دواء خافض للضغط. في معظم الأحيان، يقارن باحثو الطبية الحيوية مجتمعات إحصائية من الناس أو الحيوانات خضعت لاثنين أو أكثر من العلاجات المختلفة أو الاختبارات التشخيصية، وهم يريدون معرفة ما إذا كان هناك فرق بين استجابات المجتمعات الإحصائية التي تلقت علاجات أو اختبارات مختلفة. نستخلص في بعض الأحيان عينات متعددة من نفس المجتمع الإحصائي من الأشخاص أو الوحدات التجريبية. وهناك مثال شائع هو عندما يتم أخذ البيانات الفسيولوجية قبل وبعد علاج محدد، مثل امتصاص الأدوية أو العلاج الإلكتروني، من مجتمع إحصائي واحد من المرضى. ندعو هذا النوع من جمع البيانات في التصميم التجريبي "بالحجب". وتم مناقشة مفهوم الحجب هذا بشكل كامل في الفصل الثاني.

إن إحدى المسائل الأخرى التي كثيراً ما تكون هدف البحوث الطبية الحيوية هي ما إذا كان هناك ارتباط بين متغيرين فسيولوجيين. على سبيل المثال، هل هناك ارتباط بين بناء الجسم والوفيات؟ أو هل هناك علاقة بين امتصاص الدهون وحدوث الأورام السرطانية. أو هل هناك ارتباط بين حجم عضلة البطين في القلب وتردد نَظْم القلب غير الطبيعي؟ يتضمن هذا النوع جمع مجموعتين من البيانات وإجراء تحليل الارتباط لتحديد مقدار جودة إمكانية التنبؤ بمجتمع إحصائي من البيانات من الأخرى. وعندما نتكلم عن تحليل الارتباط، فإننا نشير إلى العلاقة الخطية بين متغيرين، والقدرة على التنبؤ بمجتمع إحصائي واحدة من البيانات عن طريق نمذجة البيانات كدالة خطي للمجتمع الإحصائي

الثاني من البيانات. ونظراً لأن تحليل الارتباط يقيس فقط العلاقة الخطية بين عمليتين أو مجموعتين من البيانات، فقد تكون العلاقات غير الخطية بين العمليتين غير واضحة. ويمكن الاطلاع على وصف أكثر تفصيلاً لتحليل الارتباط في الفصل السابع.

وأخيراً، قد يكون المهندس الطبي الحيوي، وخاصة المهندس المشارك في التصنيع، مهتماً في معرفة ما إذا كانت عملية التصنيع تحت السيطرة. قد تنشأ مثل هذه المسائل إذا كانت هناك ضوابط صارمة على مواصفات التصنيع لجهاز طبي. على سبيل المثال، إذا كان المهندس يحاول ضمان جودة إنتاج القثاطر داخل الأوعية الدموية التي يجب أن تكون بأقطار تتراوح بين 1 و 2 سنتيمتر (سم)، فقد يقوم المهندس بجمع عينات من القثاطر عشوائياً من خط التجميع على فترات عشوائية خلال النهار، وقياس أقطارها، وتحديد عدد القثاطر التي تحقق المواصفات، وتحديد ما إذا كان هناك تغيير مفاجئ في عدد القثاطر التي لا تحقق المواصفات. إذا كان مثل هذا التغيير موجودًا، فقد يبحث المهندسون عن عناصر عملية التصنيع التي تتغير مع مرور الزمن، أو المتغيرات في العوامل البيئية، أو أخطاء المُستخدِم. ويمكن للمهندس استخدام مخططات التحكم لتقييم ما إذا كانت العمليات تحت السيطرة. لم يتم تغطية طرق التحليل الإحصائي هذه في هذا الكتاب، ولكن يمكن العثور عليها في عدد من المراجع، بما في ذلك المرجع [3].

## (٣,٤) كيف نلخص البيانات بيانيًّا؟

## HOW DO WE GRAPHICALLY SUMMARIZE DATA?

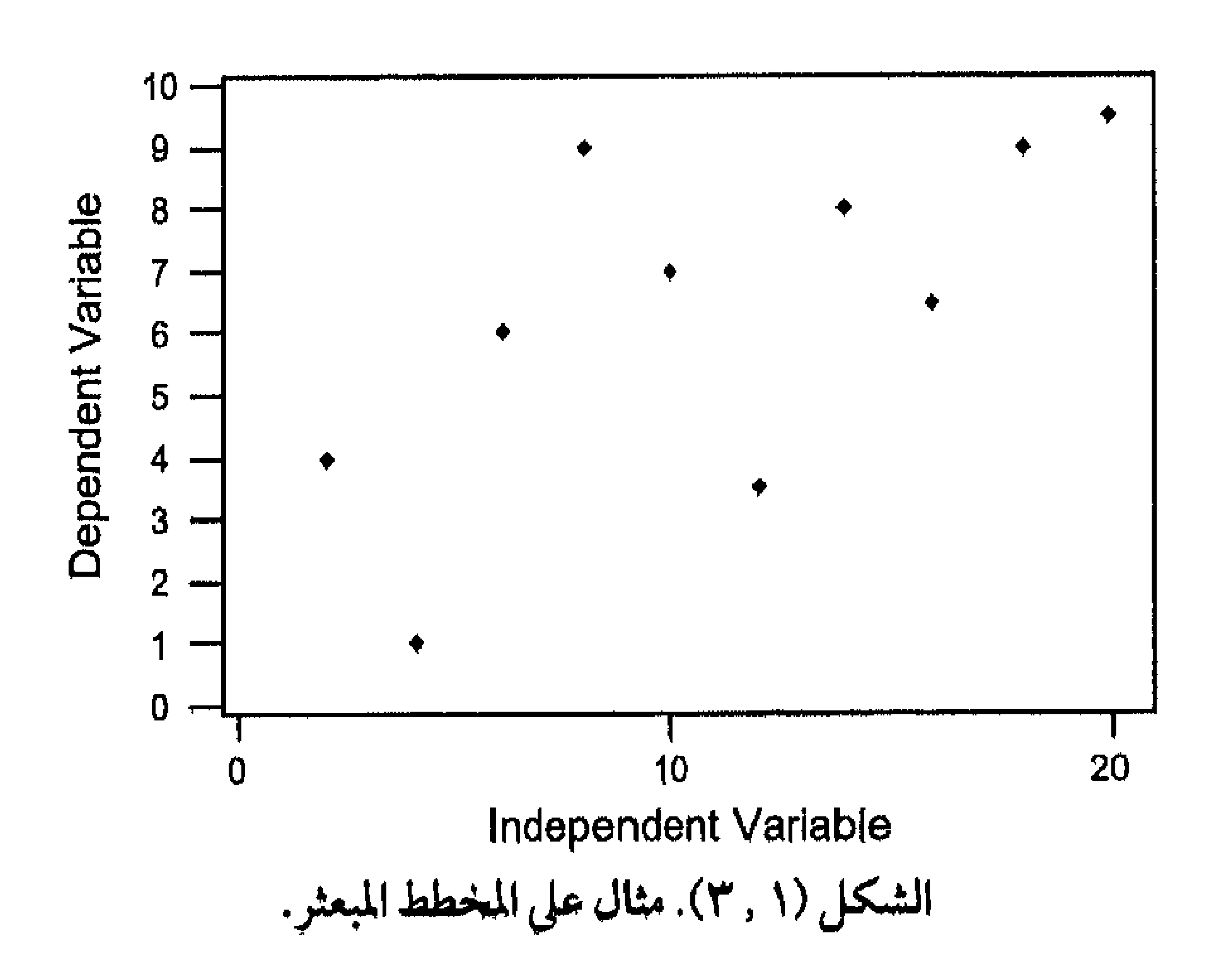
يمكننا تلخيص البيانات بشكل بياني أو عددي. و الصورة العددية هي ما نطلق عليه الإحصائيات. قبل تطبيق التحليل الإحصائي بشكل متهور، فإنه من الجيد دائماً إلقاء نظرة على البيانات الخام، عادة في شكل بياني، ومن ثم استخدام أساليب بيانية لتلخيص البيانات في شكل يسهل فيه تفسيرها.

تشمل أنواع العرض البياني الأكثر استخداماً من قبل المهندسين الطبيين الحيويين ما يلي: المخططات (الرسوم) المبعثرة، والسلاسل الزمنية، المخططات الصندوق والمؤشر، والمدرج التكراري.

إن تفاصيل إنشاء هذه الملخصات البيانية موصوفة في المراجع [3 – 6]، ولكن سوف نصفها هنا بإيجاز.

**(١ ,٤ ,٣) المخططات المبعثرة Scatterplots**

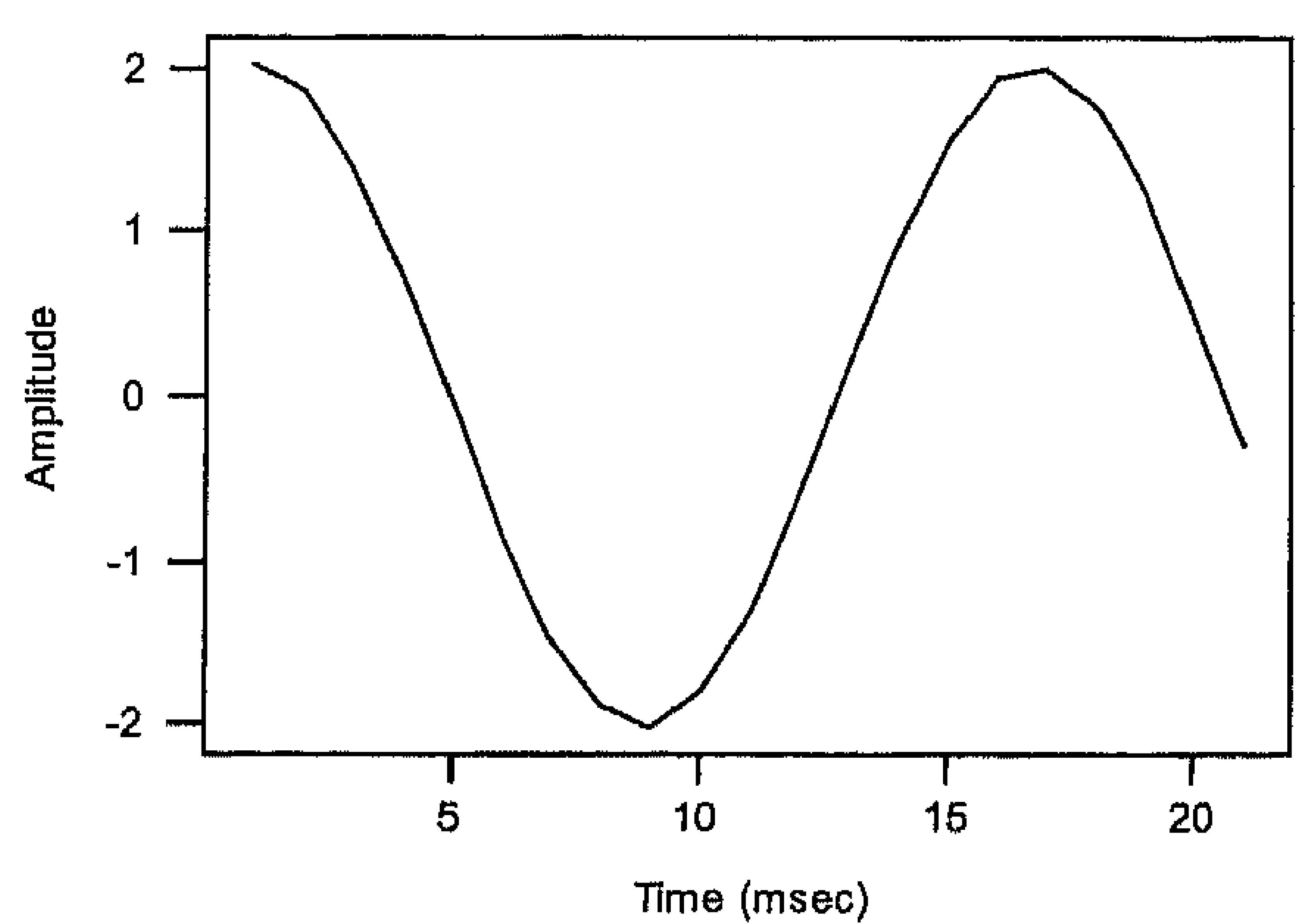
ترسم المخططات المبعثرة ببساطة الحدث لمتغير واحد بالنسبة إلى آخر. في معظم الحالات، يمكن اعتبار أحد المتغيرات المتغير المستقل (مثل الزمن أو رقم الشخص)، ويتم اعتبار المتغير الثاني المتغير التابع. يوضح الشكل (٣,١) مثالاً عن المخططات المبعثرة لمجموعتين من البيانات. بشكل عام، نحن مهتمون فيما إذا كانت هناك علاقة يمكن التنبؤ بها ترسم خريطة المتغير المستقل (مثل معدل التنفس) بالنسبة إلى المتغير الدال لدينا (مثل معدل ضربات القلب). إذا كان هناك علاقة خطية بين المتغيرين، فإنه ينبغي لنقاط البيانات أن تقع قريبة من خط مستقيم.




**الشكل (١ ,٣). مثال على المخطط المبعثر.**

### (٣,٤,٢) السلاسل الزمنية Time Series

يتم استخدام السلاسل الزمنية لرسم التغييرات في متغير كدالة في الزمن. وعادة ما يكون المتغير مقداراً فسيولوجياً يتغير مع الزمن، مثل النشاط الكهربائي في الدماغ أو تركيز الهرمون في مجرى الدم،. يوضح الشكل (٣,٢) مثالاً على مخطط سلاسل زمنية. وننظر في هذا الشكل إلى دالة الجيب البسيطة عندما يتغير مع الزمن.

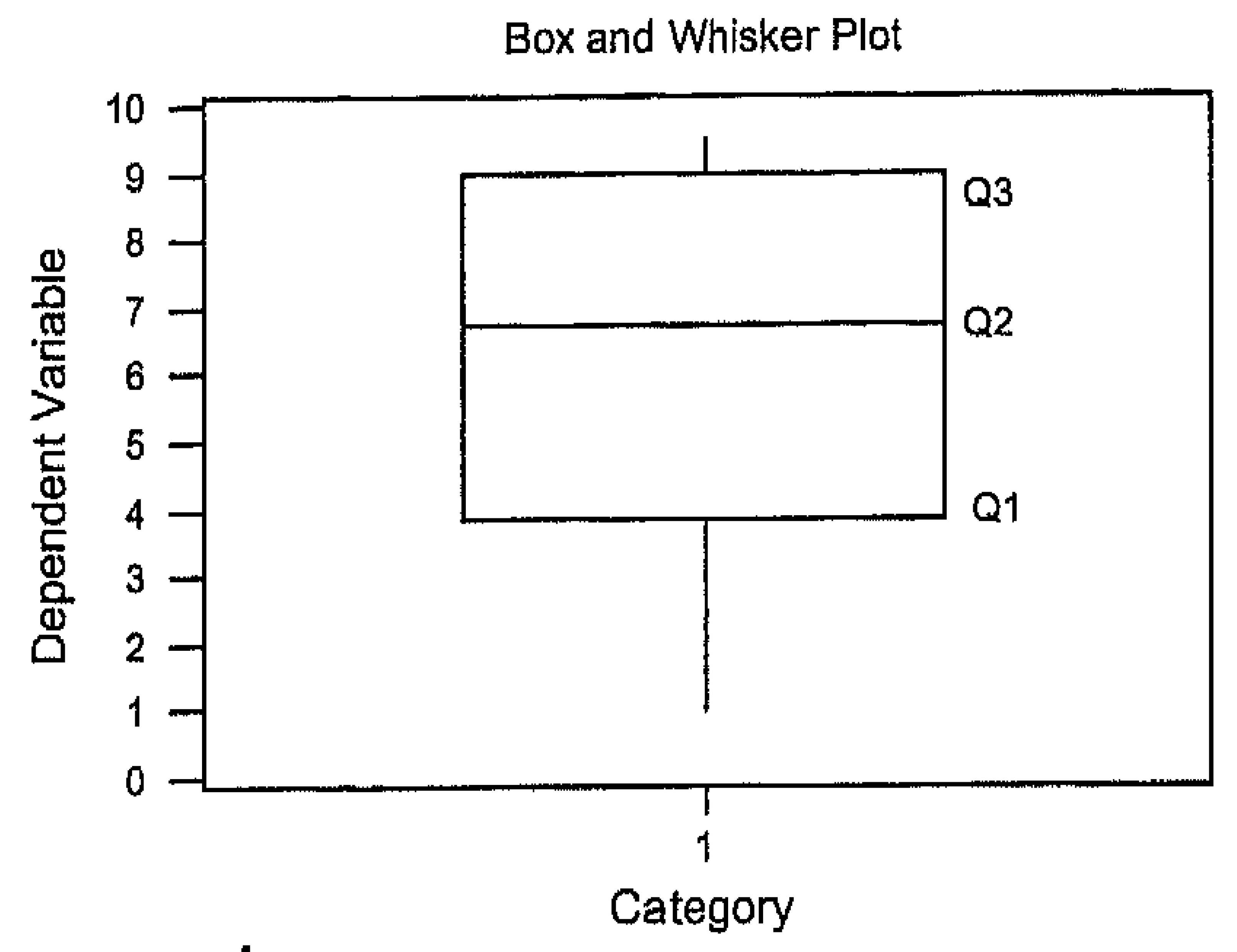



الشكل (٢ ,٣). مثال على مخطط السلاسل الزمنية. يتم رسم سعة العينات كدالة في الزمن.

### (٣,٤,٣) المخططات الصندوق والمؤشر Box-and-Whisker Plots Series

توضح هذه المخططات الشرائح الربعية (quartiles) الأولى والثانية والثالثة، بالإضافة إلى القيمة العظمى والقيمة الصغرى لقيم البيانات التي تم جمعها. إن الربع الثاني (Q2) معروف بأنه الوسيط للبيانات. هذه الكمية، على النحو المحدد في وقت لاحق في هذا الكتاب، هي نقطة البيانات الوسطى أو قيمة العينة الوسطى عندما يتم ترتيب العينات في ترتيب تنازلي. يمكن التفكير في شريحة الربيع الأول (Q1) بأنها قيمة متوسط العينات التي تقع تحت شريحة الربيع الثاني. وبالمثل، يمكن التفكير في شريحة الربيع الثالث (Q3) بأنها قيمة متوسط العينات التي تقع فوق شريحة الربيع الثاني.

تُعتبر المخططات الصندوق والمؤشر مفيدة من حيث إنها تبرز ما إذا كان هناك انحراف في البيانات أو أية قيم متطرفة غير عادية في العينات (الشكل ٣,٣).




الشكل (٣,٣). توضيح لمخطط الصندوق والمؤشر لمجتمع الإحصائي المُدرجة. ويتم تبيان شرائح الربيع الأول (Q1) والثاني (Q2) والثالث (Q3). بالإضافة إلى ذلك، تمتد الشعيرات إلى القيمة العظمى والقيمة الصغرى لمجتمع العينة.

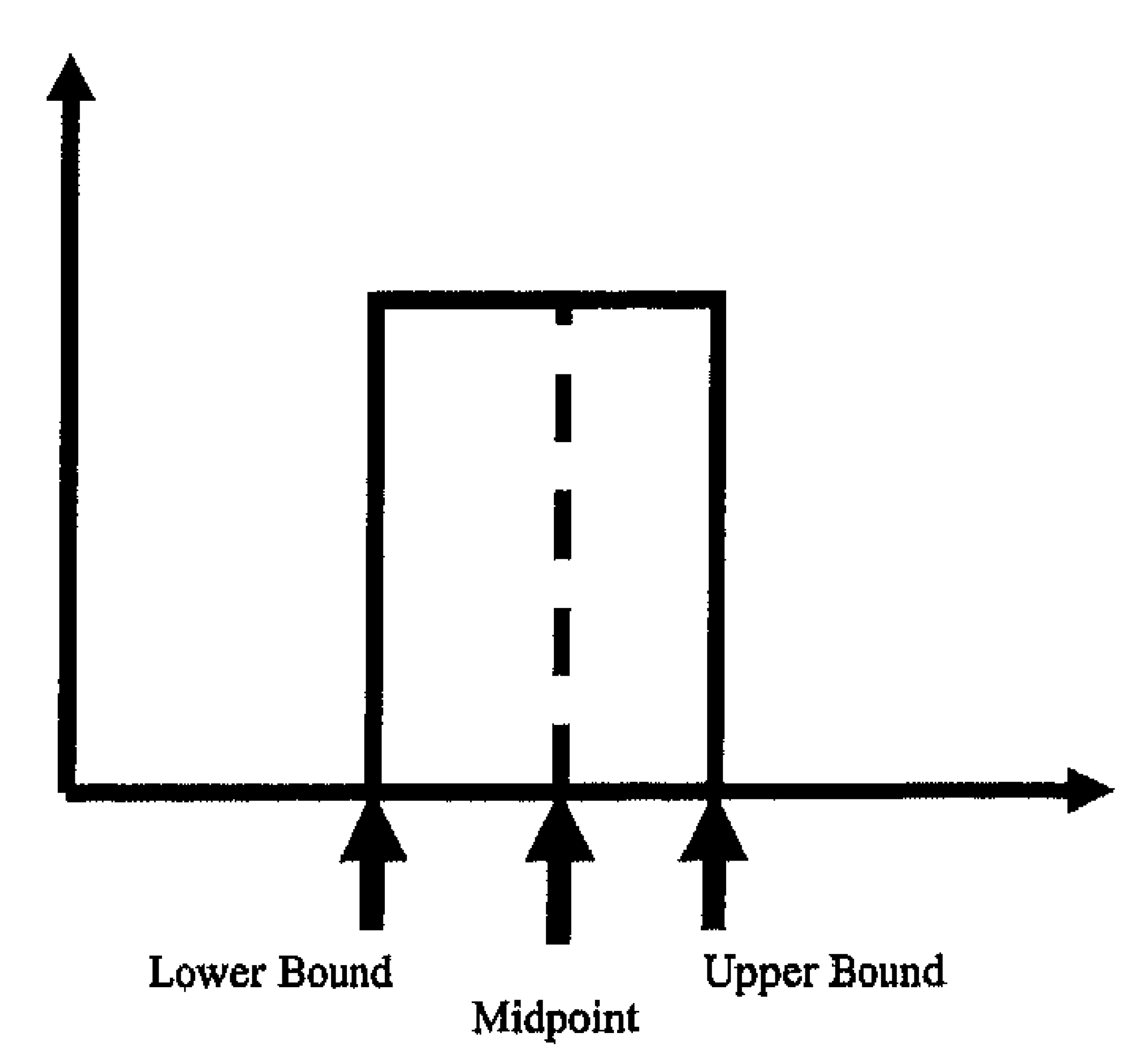
### (٣,٤,٤) المدرج التكراري Histogram

يتم تعريف المدرج التكراري بأنه توزيع الترددات. عند الأخذ في الاعتبار $N$ عينات أو قياسات ، $x_i$ ، وفى المدى من $X_{\min}$ إلى $X_{\max}$ ، فإنه يتم تجميع العينات في فترات زمنية غير متداخلة (صناديق)، وعادة متساوية العرض (الشكل ٣,٤). يكون عدد الصناديق عادة من 7 – 14 ، وهذا متوقف على طبيعة البيانات. وبالإضافة إلى ذلك ، نتوقع عادة أن يكون لدينا ثلاث عينات على الأقل لكل صندوق [7]. ويمكن أيضاً استخدام قاعدة Sturgess (سترجس) [6] لتقدير عدد الصناديق والتي تُعطى بالعلاقة التالية :

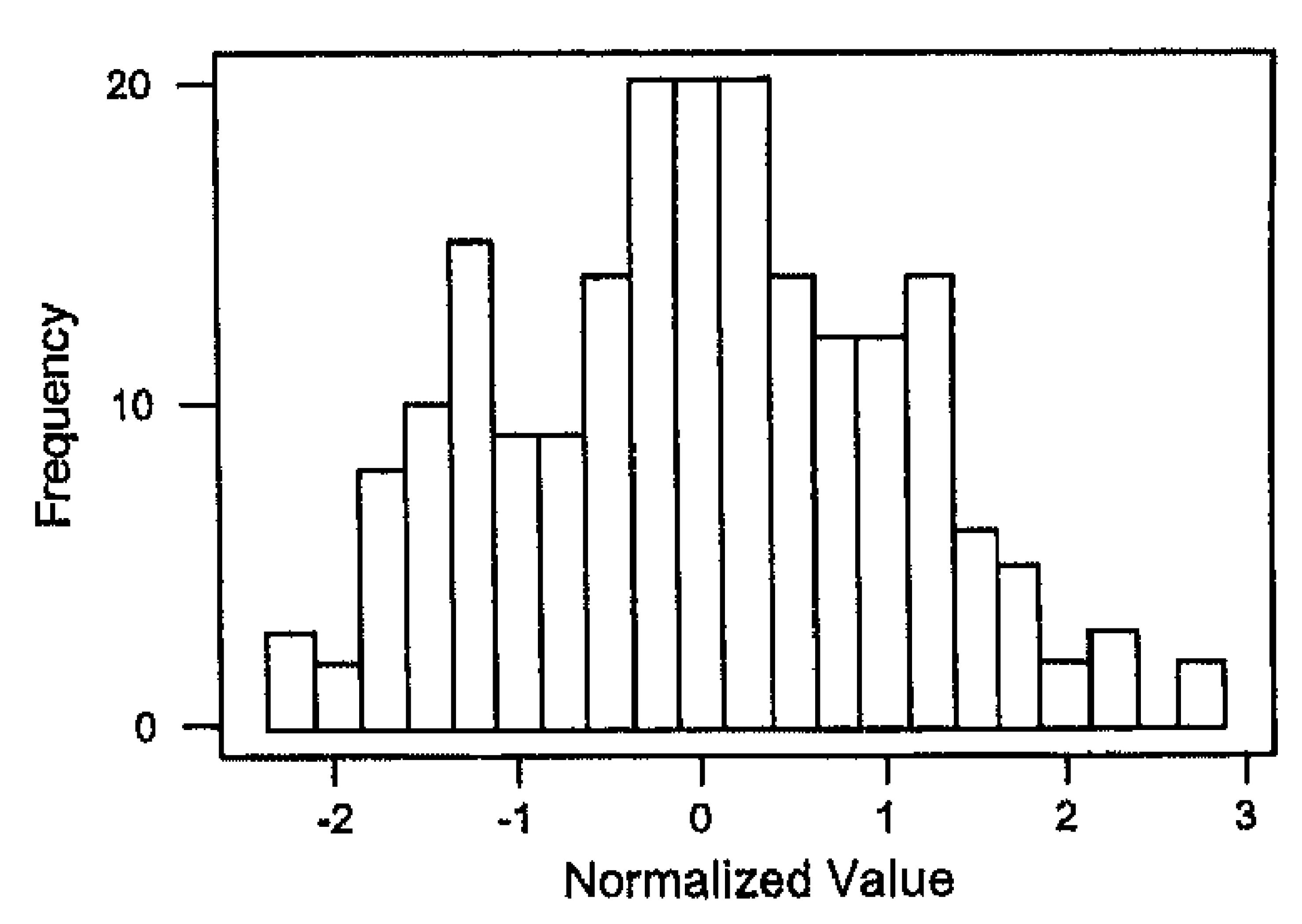
$$k = 1 + 3.3 \log (n)$$

حيث $k$ هو عدد الصناديق و $n$ هو عدد العينات.

كل صندوق من المدرج التكراري لديه حد أدنى ، وحد أعلى ، ونقطة منتصف. يتم إنشاء المدرج التكراري من خلال رسم عدد العينات في كل صندوق. يوضح الشكل (٣،٥) المدرج التكراري لـ 1000 عينة مأخوذة من توزيع عادي له متوسط $(\mu)$=0 وانحراف معياري 1.0=$(\sigma)$. لدينا على المحور الأفقي قيمة العينة وعلى المحور العمودي عدد حدوث العينات التي تقع ضمن الصندوق.




الشكل (٣ ,٤). أحد صناديق مخطط المدرج التكراري يتم تعريف الصندوق من خلال حد أدنى، ونقطة منتصف، وحد أعلى.




الشكل (٣ ,٥). مثال لمخطط مدرج تكراري يتم رسم قيمة المقدار أو العينة على المحور الأفقي في حين يتم رسم تردد حدوث ذلك المقدار أو العينة على طول المحور العمودي.

يوجد مقداران مفيدان في وصف المدرج التكراري هما التردد المطلق والتردد النسبي في واحد أو أكثر من الصناديق. يتم تعريف هذه الكميات كما يلي :

أ) $f_i$= التردد المطلق في الصندوق رقم $i$ ؛

ب) $f_i/n$= التردد النسبي في الصندوق رقم $i$، حيث n هي العدد الإجمالي للعينات والذي يجري تلخيصه في المدرج التكراري.

هناك عدد من الخوارزميات التي تستخدمها الأجهزة الطبية الحيوية لتشخيص أو كشف التشوهات في وظيفة بيولوجية ؛ وتستفيد هذه الخوارزميات من المدرج التكراري البيانات التي تم جمعها والترددات النسبية المرتبطة بالصناديق المُختارة [8]. وفي كثير من الأحيان يمكن التمييز بين الوظائف الفسيولوجية الطبيعية وغير الطبيعية (أصوات التنفس، وتغيّر معدل ضربات القلب، والمحتوى الترددي للإشارات الفسيولوجية الكهربائية) من خلال مقارنة الترددات النسبية في الصناديق المُستهدفة من المدرج التكراري للبيانات التي تمثل هذه العمليات البيولوجية.

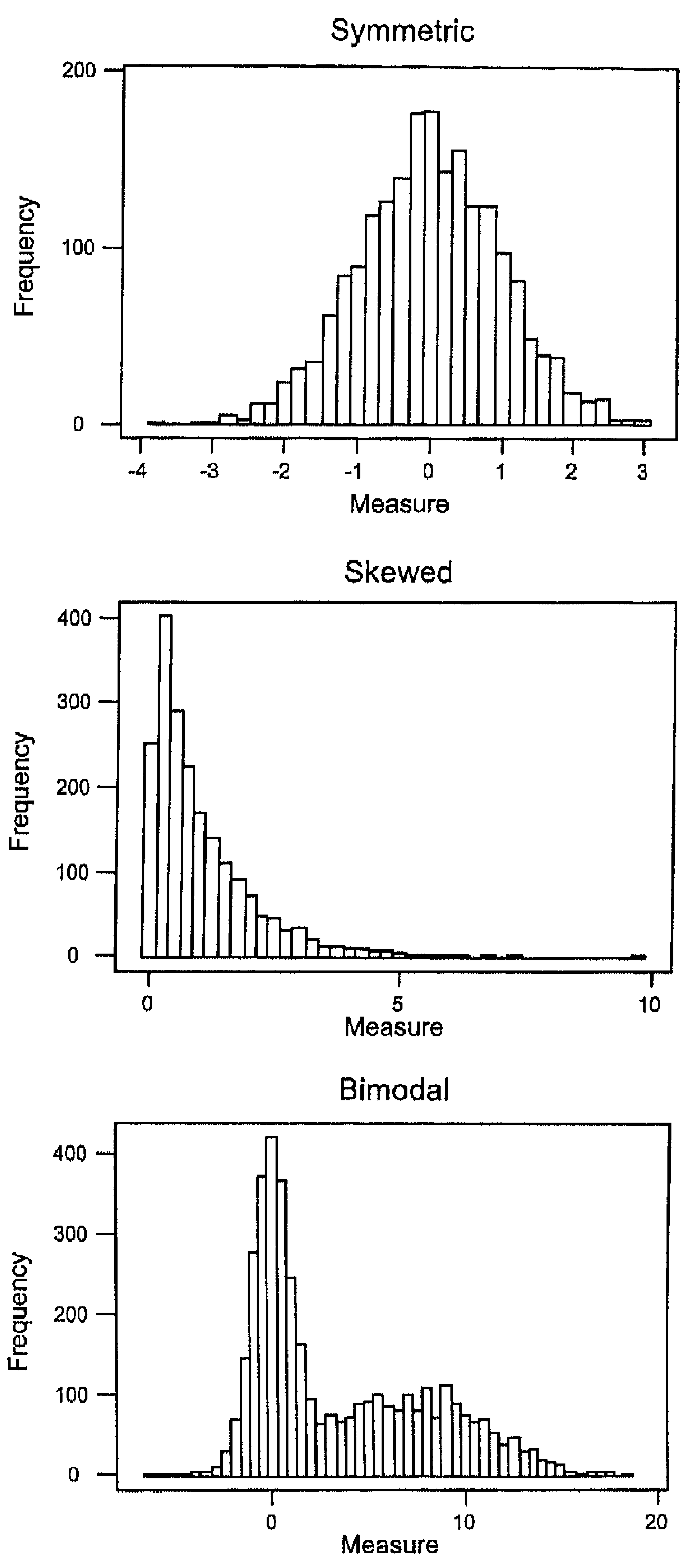
يمكن أن يأخذ المدرج التكراري عدة أشكال. ويُشار إلى الأشكال، المُوضحة في الشكل (٣.٦)، بالمتناظرة، أو غير المتماثلة (المنحرفة)، أو ثنائية الشكل (الدارج).

ويمكن أن يُعزى المدرج التكراري غير المتماثل إلى ما يلي [9]:

١- الآليات ذات الاهتمام التي تولّد البيانات (على سبيل المثال، الآليات الفسيولوجية التي تحدد الفترات الزمنية من ضربة إلى ضربة في القلب).

٢- تشويش صنعي بسبب عملية القياس أو تغيير في الآلية الأساسية مع مرور الزمن (على سبيل المثال، قد يكون هناك تبدلات متغيرة مع الزمن في عملية التصنيع التي تؤدي إلى تغيير في إحصائيات عملية التصنيع مع مرور الزمن).

٣- مزج للمجتمعات الإحصائية التي يتم أخذ العينات منها (وهذا نموذجياً مصدر المدرج التكراري ثنائي الشكل).




الشكل (٦ ,٣). أمثلة على المدرج التكراري المتناظر (أعلى)، وغير المتماثل (الوسط)، وثنائي الشكل (أسفل). في كل حالة، تم أخذ 2000 عينة من المجتمعات الإحصائية.

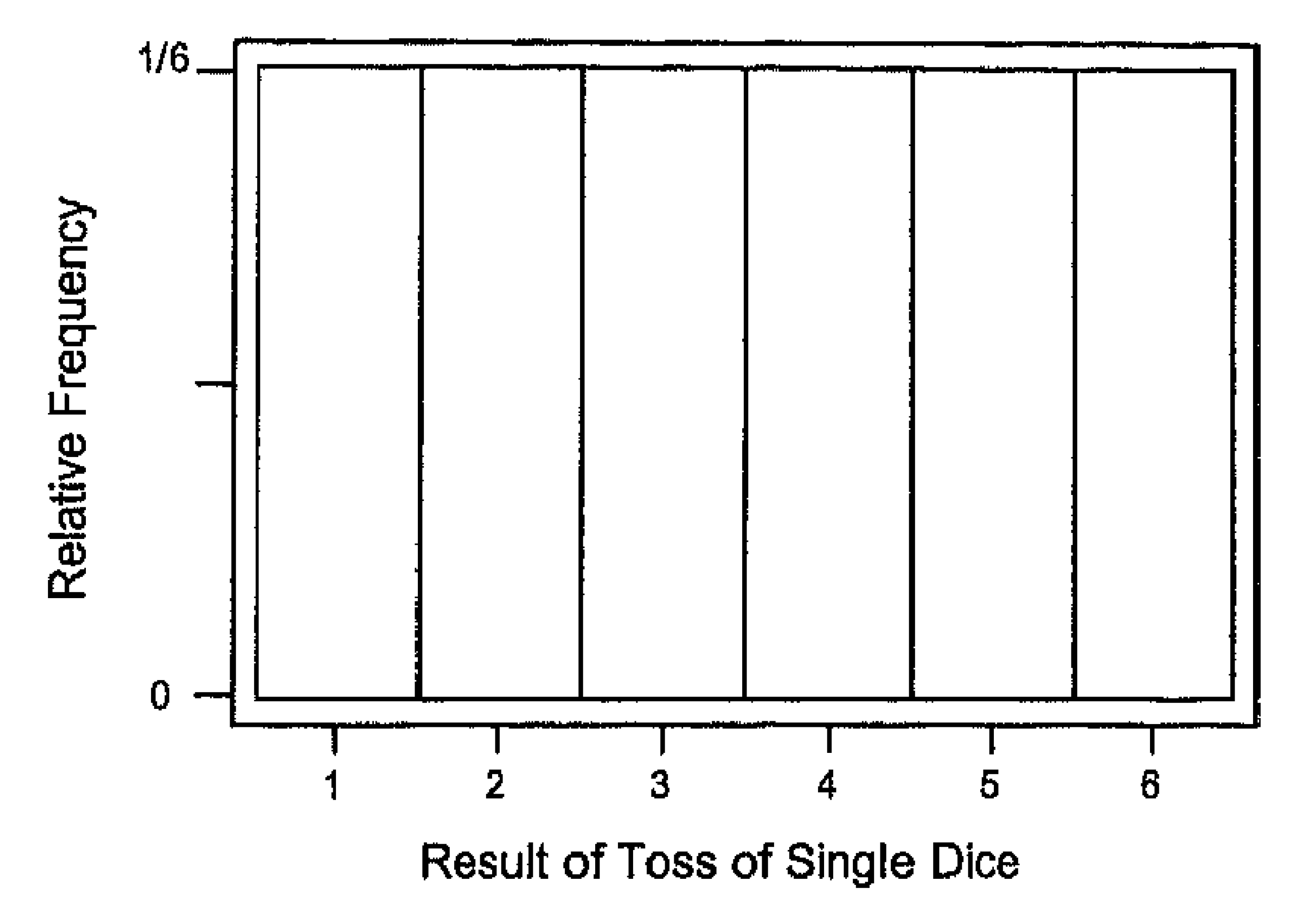
يُعد المدرج التكراري مهمًّا لأنه بمثابة تقدير تقريبي لدالة كثافة الاحتمال الحقيقية أو توزيع الاحتمال للعملية العشوائية الأساسية التي يجري جمع العينات منها.

إن دالة كثافة الاحتمال أو توزيع الاحتمال هي دالة تحدد مقدار احتمال حدوث حادثة عشوائية ، $x_i$ . عندما يكون الحدث العشوائي الأساسي متقطعاً في الطبيعة ، فإننا نشير إلى دالة كثافة الاحتمال بأنها دالة كتلة الاحتمال [10]. وفي كلتا الحالتين، تصف الدالة الطبيعة الاحتمالية للمتغير أو الحدث العشوائي الأساسي ، ويسمح لنا أن نتوقع احتمال ملاحظة نتيجة محددة ، $x$ (مُمثَّلة بمتغير عشوائي) ، لتجربة ما. إن دالة التوزيع التراكمي هي ببساطة مجموع الاحتمالات لمجتمع إحصائي من النتائج ، حيث تكون النتيجة أقل من أو تساوى قيمة محددة لـ $x$ .

دعونا نأخذ في الاعتبار المتغير العشوائي الذي من أجله تكون دالة كثافة الاحتمال مُعرَّفة بشكل جيد (فيما يتعلق بمعظم ظواهر العالم الحقيقي ، فإن مثل نماذج الاحتمال هذه تكون غير معروفة). إن المتغير العشوائي هو النتيجة لرمية واحدة للنرد (الزهر). ليكن لدينا نرد واحد واضح ذو ستة أوجه ، فإن احتمال تدحرج الرقم ستة من رمية النرد هو 1 من 6. وفي الواقع ، فإن احتمال رمي الرقم واحد هو أيضاً 1 من 6. إذا أخذنا بعين الاعتبار جميع النتائج المحتملة لرميات النرد ورسم احتمال ملاحظة أي واحدة من تلك النتائج الستة في رمية واحدة ، سيكون لدينا مخطط مثل ذلك المُبيّن في الشكل (٣.٧).

يبين هذا المخطط دالة كثافة الاحتمال أو دالة كتلة الاحتمال لرمية من النرد. ويُعرف هذا النوع من نماذج الاحتمال بالتوزيع المنتظم لأن كل نتيجة لديها نفس احتمال الحدوث بالضبط (1/6 في هذه الحالة).

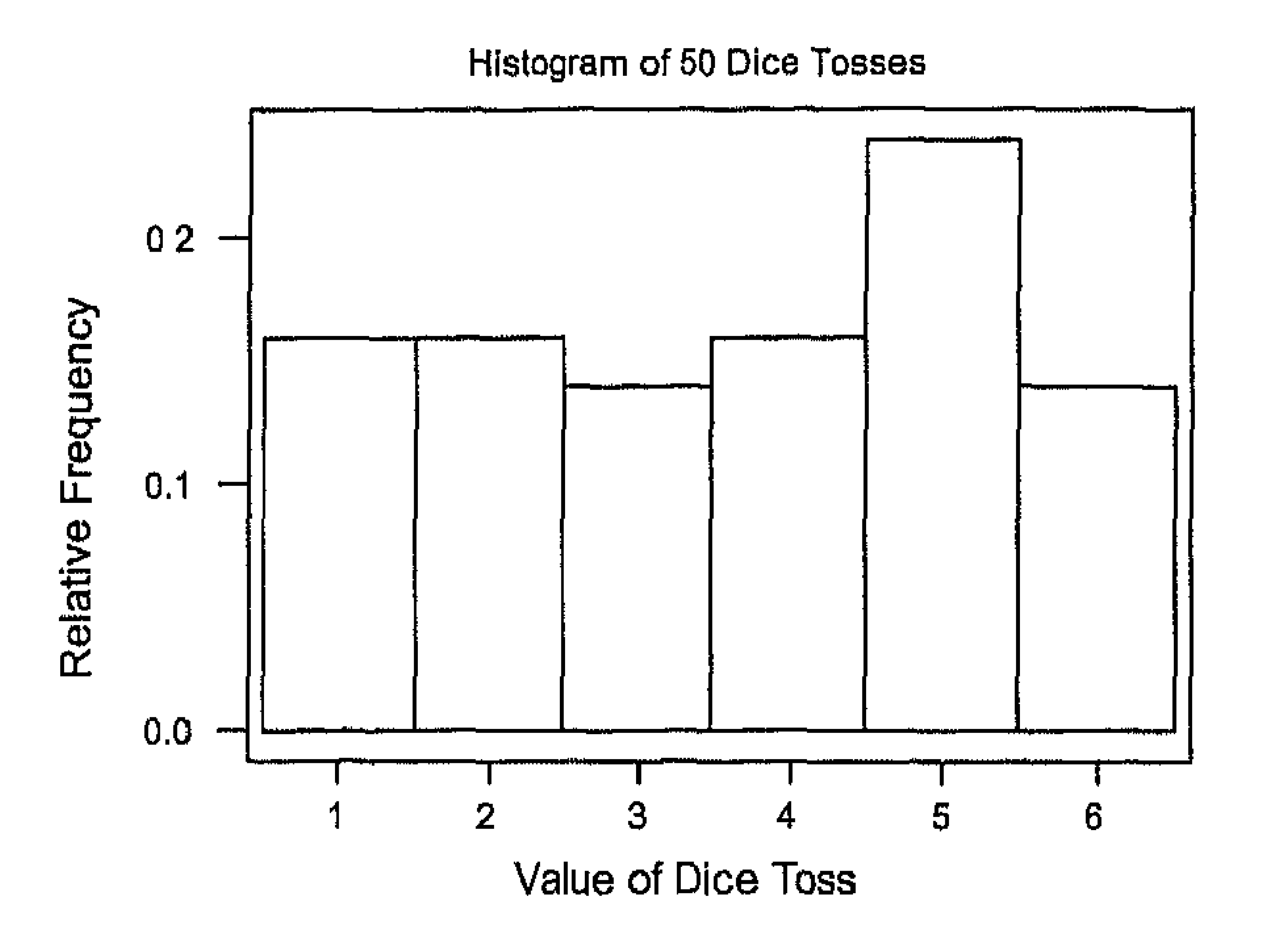


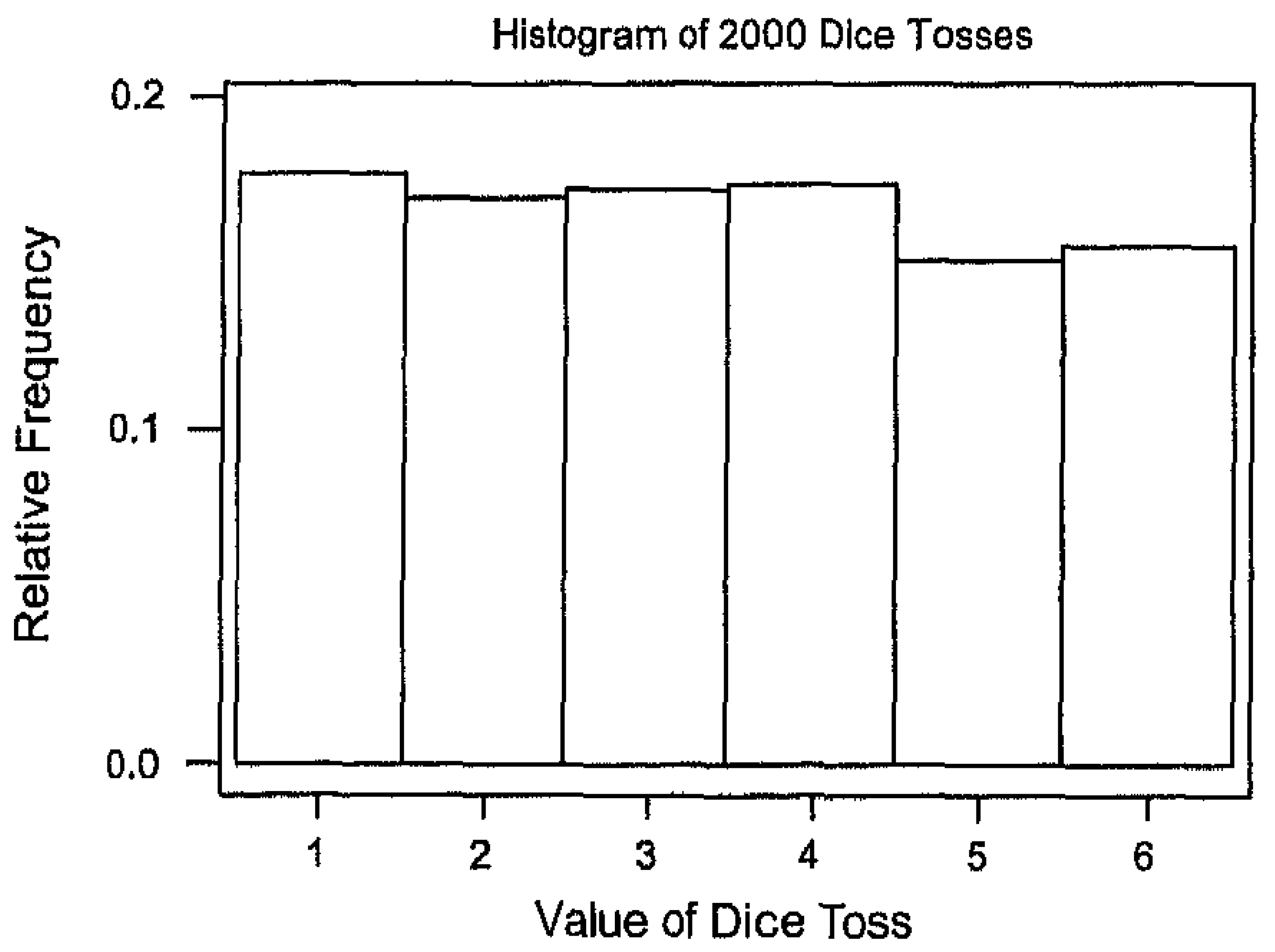




**الشكل (٣،٧). دالة كثافة الاحتمال لمتغير عشوائي متقطع (دالة كتلة الاحتمال). في هذه الحالة، فإن المتغير العشوائي هو قيمة الرمية لنرد (زهر) واحد. لاحظ أن كل نتيجة من النتائج الست المحتملة لديها احتمال حدوث مقداره 1 من 6. إن دالة كثافة الاحتمال هذا معروفة أيضاً بأنها توزيع الاحتمال المنتظم.**

نحن نعلم توزيع الاحتمال الحقيقي بالنسبة لرمية النرد. ومع ذلك، بالنسبة لمعظم العمليات العشوائية في العالم الحقيقي، وخاصة العمليات البيولوجية، فإننا لا نعرف كيف تبدو دالة كثافة الاحتمال الحقيقي أو دالة كتلة الاحتمال. ونتيجة لذلك، علينا استخدام المدرج التكراري، الذي يتم إنشاؤه من عينة صغيرة، في محاولة لتقدير التوزيع الأفضل للاحتمال أو نموذج الاحتمال لوصف الظاهرة في العالم الحقيقي. إذا عدنا إلى مثال رمية النرد، فإنه يمكننا في الواقع رمي النرد عدة مرات ورؤية مقدار تطابق المدرج التكراري، الذي تم الحصول عليه من البيانات التجريبية، مع دالة كتلة الاحتمال الحقيقي للنرد المثالي سداسي الوجوه. يوضح الشكل (٣،٨) المدرجات التكرارية لنتائج رميات عددها 50 و 1000 رمية لنرد واحد. لاحظ أنه حتى مع 50 رمية أو عينة، فإنه من الصعب تحديد كيف يمكن أن يبدو توزيع الاحتمال الحقيقي. ومع ذلك، عندما نقترب من 1000 عينة، فإن المدرج التكراري يقترب من دالة كتلة الاحتمال الحقيقي (توزيع منتظم) لرمية النرد. ولكن، لا يزال هناك بعض التباين من صندوق إلى صندوق لذلك لا يبدو منتظماً مثل توزيع الاحتمال

المثالي المُوضح في الشكل (٣.٧). إن الرسالة التي يمكن قراءتها من هذا التوضيح هي أن معظم الأبحاث الطبية الحيوية تنشر النتائج لعدد قليل من العينات. ويتضح من مثال النرد أنه من الصعب جدًّا تحديد الإحصائيات للعملية العشوائية الأساسية من عينة صغيرة، ولكن معظم الأبحاث الطبية الحيوية تعتمد على بيانات من عينات صغيرة.




**الشكل (٨ , ٣). المدرجات التكرارية تمثل نتائج التجارب التي يتم فيها رمي نرد واحد 50 مرة (أعلى) و 2000 مرة (أسفل)، على التوالي لاحظ أنه كلما ازداد حجم العينة، اقترب المدرج التكراري من توزيع الاحتمال الحقيقي المُوضح في الشكل (٣.٧).**

## (٣,٥) النهج العام للتحليل الإحصائي
## GENERAL APPROACH TO STATISTICAL ANALYSIS

لقد جمعنا بياناتنا الآن وألقينا النظر على بعض الملخصات البيانية للبيانات. والآن سوف نستخدم ملخصاً عدديًا، معروفاً أيضاً باسم الإحصائيات، في محاولة لوصف طبيعة المجتمع الإحصائي أو العملية التي أخذنا منها العينات. نفترض، من هذه الإحصائيات الوصفية، نموذج احتمال أو توزيع احتمال للمجتمع الإحصائي أو العملية ومن ثم نختار الاختبارات الإحصائية المناسبة لاختبار الفرضيات أو اتخاذ القرارات. ومن المهم ملاحظة أن الاستنتاجات التي يمكن للمرء استخلاصها من تجربة إحصائية تعتمد على مدى ملاءمة نموذج الاحتمال المُفترض للمجتمع الإحصائي أو العملية.

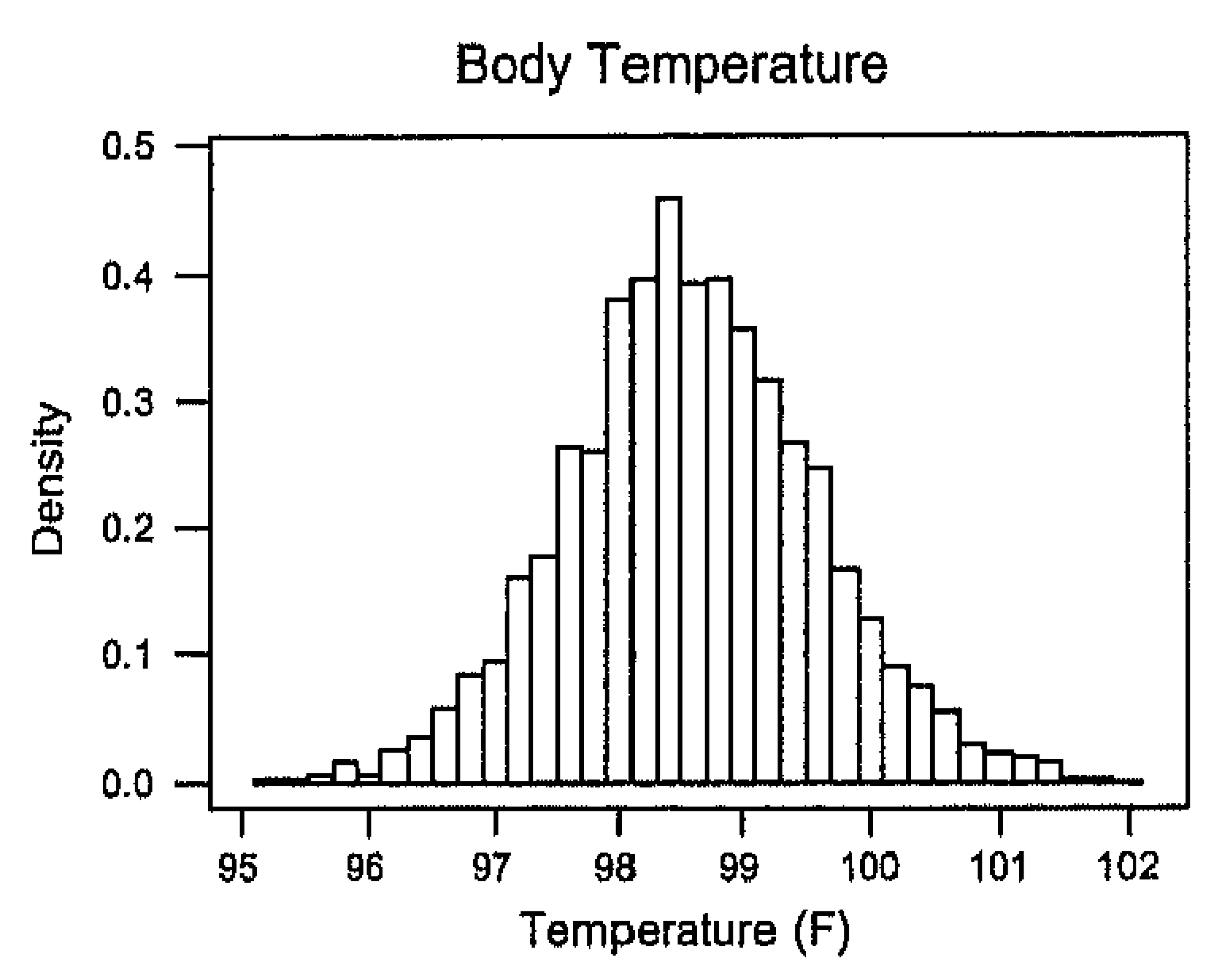
وكما جاء في المقدمة، يحاول مهندسو الطبية الحيوية اتخاذ قرارات حول المجتمعات الإحصائية أو العمليات التي لديهم وصول محدود إليها. وهكذا، فهم يصممون التجارب ويجمعون العينات التي يعتقدون أنها إلى حد ما تمثل المجتمع الإحصائي أو العملية. وبغض النظر عن نوع التحليل الإحصائي الذي سوف ينتج عن البحث أو الدراسة، فإنه ينبغي لجميع أنواع التحليل الإحصائي أن تتبع نفس النهج العام:

١ - قياس عدد محدود من عينات تمثيلية من المجتمع الإحصائي الأكبر.

٢ - تقدير الإحصاءات الحقيقية للمجتمع الإحصائي الأكبر من إحصائيات العينة.

يجب هنا معالجة بعض المفاهيم الهامة. المفهوم الأول واضح إلى حد ما. غالباً ما يكون مستحيلاً أو غير عملي أخذ القياسات أو الملاحظات من المجتمع الإحصائي الكامل. وهكذا، فإن المهندس الطبي الحيوي عادة ما يختار عينة أصغر وأكثر عملية تمثل المجتمع الإحصائي ومدى المتغير في المجتمع الإحصائي الأكبر. على سبيل المثال، قد لا نستطيع قياس درجة حرارة الجسم في حالة الراحة لكل إنسان على وجه الأرض للحصول على تقدير لدرجة حرارة الجسم الطبيعية والمجال الطبيعي لها. يتركز اهتمامنا هنا على معرفة ما هي درجة حرارة الجسم الطبيعية، في المتوسط،

للإنسان السليم والمجال الطبيعي لدرجة الحرارة في حالة الراحة بالإضافة إلى احتمال أو إمكانية قياس درجة حرارة الجسم الخاصة في الحالات الصحية، وفي ظروف الراحة. في محاولة لتحديد الخصائص أو نموذج الاحتمال الأساسي لدرجة حرارة الجسم للأفراد الأصحاء، وفي حالة الراحة، سيختار الباحث، على نحو عشوائي، عينة من الأفراد الأصحاء وفي حالة الراحة ويقيس درجة حرارتهم الخاصة في حالة الراحة بواسطة مقياس حرارة. يجب أن يراعي الباحثون تكوين وحجم مجتمع العينة لتمثيل المتغير بشكل كاف في المجتمع الإحصائي. ويجب على الباحث تحديد ما يميّز الفرد الطبيعي والسليم، مثل العمر والحجم والعرق والجنس والصفات الأخرى. إذا كان الباحث يجمع بيانات درجة حرارة الجسم من عينة مؤلفة من 3000 شخص مثلاً، فقد يرسم المدرج تكرارياً لدرجات الحرارة التي تم قياسها من الـ 3000 شخص ويخلص في النهاية إلى المدرج التكراري التالي المُبين في الشكل (٣،٩). وقد يحسب الباحث أيضاً بعض الإحصائيات الوصفية الأساسية للـ 3000 عينة، مثل متوسط العينة (الوسيط الحسابي) (average (mean)، الوسيط (median) (القيمة الواقعة في المنتصف)، والانحراف المعياري (standard deviation).




**الشكل (٩، ٣). المدرج التكراري لـ 2000 قيمة لدرجة حرارة الجسم الداخلية تم جمعها من المجتمع الإحصائي مُوزَّعة بشكل طبيعي.**

يحاول الباحث، بعد تقدير إحصائيات العينة من مجتمع إحصائي، استخلاص استنتاجات حول المجتمع الإحصائي الأكبر (الحقيقية). إن السؤال الأكثر أهمية الذي يجب طرحه عند مراجعة الإحصائيات والاستنتاجات المُستخلصة من مجتمع إحصائي هو مقدار جودة تمثيل مجتمع العينة للمجتمع الإحصائي الأكبر.

نستخدم، بعد جمع البيانات، بعض الإحصائيات الوصفية الأساسية لتلخيص البيانات. وتشمل هذه الإحصائيات الوصفية الأساسية المقادير العامة التالية: النزعة المركزية، والمتغير، والارتباط.

## (٣,٦) الإحصاء الوصفي
## DESCRIPTIVE STATISTICS

هناك عدد من الإحصائيات الوصفية التي تساعدنا على تصوّر توزيع المجتمع الإحصائي. ويعبارة أخرى، إن هدفنا النهائي هو افتراض نموذج احتمال أساسي للمجتمع الإحصائي ومن ثم اختيار التحليلات الإحصائية المناسبة لنموذج الاحتمال هذا.

عندما نحاول استخلاص استنتاجات حول المجتمع الإحصائي الأكبر أو العملية من العينة الأصغر للبيانات لدينا، فإننا نفترض أن النموذج الأساسي لأي عينة، أو "حدث"، أو مقدار (نتيجة التجربة) هو كما يلي:

$X = \mu$ ± الفروق الفردية ± العوامل الظرفية ± المتغيرات غير المعروفة.

حيث $X$ هو المقدار أو قيمة العينة لدينا ويتأثر بـ $\mu$، الذي هو متوسط المجتمع الإحصائي الحقيقي؛ الفروق الفردية مثل التركيب الوراثي، والتدريب، والدافع، والحالة الفيزيائية؛ والعوامل الظرفية مثل العوامل البيئية؛ والمتغيرات غير المعروفة مثل العوامل غير المعروفة / العوامل غير المُقاسة التي تتصرف بطريقة لا يمكن التنبؤ بها من لحظة إلى لحظة.

وبعبارة أخرى، عندما نقوم بإجراء قياس أو ملاحظة، فإن القيمة المُقاسة تمثل أو تكون متأثرة ليس فقط بإحصائيات المجتمع الإحصائي، مثل متوسط المجتمع الإحصائي، ولكن بعوامل مثل المتغير البيولوجي من فرد إلى فرد، والعوامل البيئية (الزمن، ودرجة الحرارة، والرطوبة، والإضاءة، والأدوية، وما إلى ذلك)، والعوامل العشوائية التي لا يمكن التنبؤ بها تماماً من لحظة إلى لحظة. تؤدي جميع هذه العوامل إلى إنشاء المدرج التكراري لبيانات العينة، التي قد تعكس أو لا تعكس دالة كثافة الاحتمال الحقيقي للمجتمع الإحصائي. إذا قمنا بعمل جيد في التصميم التجريبي وجمعنا عددًا كافيًا من العينات، فإنه ينبغي للمدرج التكراري والإحصائيات الوصفية لمجتمع العينة أن تعكس بشكل متقارب دالة كثافة الاحتمال الحقيقي والإحصائيات الوصفية للمجتمع الإحصائي الحقيقي. إذا كان هذا هو الحال، يمكننا بعد ذلك القيام باستنتاجات حول المجتمع الإحصائي الأكبر من مجتمع العينة الأصغر. وإذا كان مجتمع العينة لا يعكس تغيّر المجتمع الإحصائي الحقيقي، فقد تكون عندئذ الاستنتاجات التي قمنا باستخلاصها من التحليل الإحصائي لبيانات العينة ذات قيمة صغيرة.

هناك عدد من نماذج الاحتمال المفيدة لوصف العمليات البيولوجية وعمليات التصنيع. وتشمل التوزيع الطبيعي، وتوزيع بواسون، وتوزيع الأسس، وتوزيع جاما [10]. سنركز في هذا الكتاب على المجتمعات الإحصائية التي تتبع توزيعاً طبيعياً لأن هذا هو توزيع الاحتمال، الأكثر تكراراً والذي تتم مصادفته، المُستخدم في وصف المجتمع الإحصائي. وعلاوة على ذلك، تفترض الطرق الأكثر استخداماً للتحليل الإحصائي أن تكون البيانات مُنمذجة جيداً بواسطة التوزيع الطبيعي ("منحنى الجرس"). ومن المهم ملاحظة أن العديد من العمليات البيولوجية غير مُنمذجة بشكل جيد بواسطة التوزيع الطبيعي (مثل تغيّر معدل ضربات القلب)، والإحصائيات المرتبطة بالتوزيع

الطبيعي ليست مناسبة لمثل هذه العمليات. في مثل هذه الحالات ، قد تخدم الإحصائيات غير البارامترية (nonparametric)، التي لا تفترض نوعاً محدداً لتوزيع البيانات ، الباحث بشكل أفضل في فهم العمليات واتخاذ القرارات. ومع ذلك، فإن استخدام التوزيع الطبيعي والإحصائيات المرتبطة به غالباً ما يكون كافياً باعتبار نظرية النزعة المركزية، التي تنص ببساطة على أن جمع العمليات العشوائية ذات التوزيع العشوائي سينتج عنه متغير عشوائي ذو توزيع طبيعي. يمكن للمرء الافتراض أن معظم الظواهر البيولوجية تنتج من جمع العمليات العشوائية.

**(٣,٦,١) مقاييس النزعة المركزية Measures of Central Tendency**

هناك العديد من المقاييس التي تعكس النزعة المركزية أو التركيز لمجتمع العينة : متوسط العينة (المتوسط الحسابي)، وسيط (منتصف) العينة ، ومنوال العينة.

يمكن تقدير متوسط العينة من مجتمع من العينات ، $x_i$ ، حيث $i$ هو رقم العينة ، باستخدام الصيغة أدناه.

مُعطى $n$ نقطة من البيانات ، $x_1, x_2, \ldots, x_n$ :

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

في الواقع نحن لا نعرف عادة المتوسط الحقيقي ، μ، للمجتمع الإحصائي، بدلاً من ذلك نحاول تقدير المتوسط الحقيقي ، μ، للمجتمع الإحصائي الأكبر. عندما يصبح حجم العينة كبيراً، ينبغي لمتوسط العينة، $\overline{x}$ ، أن يقترب من المتوسط الحقيقي ، μ، على افتراض أن إحصائيات المجتمع الإحصائي أو العملية لا تتغير بمرور الزمن أو بتغير المكان.

إن إحدى المشاكل التي تنشأ عند استخدام متوسط العينة لتمثيل النزعة المركزية للمجتمع الإحصائي هي أن متوسط العينة عرضة للقيم المتطرفة. يمكن أن يكون هذا مشكلة وغالباً مضللاً عند وصف متوسط المجتمع الإحصائي المنحرف بشكل كبير.

على سبيل المثال، عند وصف الدخل لمجتمع من خريجي الجامعات الجدد التي تضم لاعب كرة سلة (NBA) وقع عقداً بعدة ملايين من الدولارات، فإن متوسط الدخل المُقدّر سوف يكون أكبر بكثير مما يكسب معظم الخريجين. إن التحريف (سوء التمثيل) نفسه واضح في كثير من الأحيان عند وصف قيمة المتوسط للمنازل في منطقة جغرافية محددة حيث يمكن لعدد قليل من المنازل قيمتها في حدود المليون إخفاء حقيقة أن عدة مئات من المنازل الأخرى قيمتها أقل من 200000 دولار أمريكي.

ثمة مقياس آخر مفيد لتلخيص النزعة المركزية للمجتمع الإحصائي هو وسيط العينة. إن قيمة الوسيط لمجتمع إحصائي من الملاحظات أو العينات، $x_i$ ، هو الملاحظة الوسطى عندما يتم ترتيب العينات، $x_i$ ، في ترتيب تنازلي.

على سبيل المثال، إذا كان لدينا القيم التالية للحجم المدّي للرئة:

2, 1.5, 1.3, 1.8, 2.2, 2.5, 1.4, 1.3,

فإنه يمكننا إيجاد قيمة الوسيط أولاً من خلال ترتيب البيانات ترتيبًا تنازليًا:

2.5, 2.2, 2.0, 1.8, 1.5, 1.4, 1.3, 1.3,

بعد ذلك نشطب القيم على كل نهاية حتى نصل إلى قيمة المنتصف:

~~2.5~~, ~~2.2~~, ~~2.0~~, 1.8, 1.5, ~~1.4~~, ~~1.3~~, ~~1.3~~,

في هذه الحالة، هناك قيمتان للوسيط، وهكذا، فإن الوسيط هو متوسط كلتا القيمتين، وهو 1.65.

لاحظ أنه إذا كان عدد العينات، $n$، عددًا فرديًا، فإن الوسيط سيكون الملاحظة الوسطى. وإذا كان حجم العينة، $n$، عددًا زوجيًا، فإن الوسيط عندئذ يساوي متوسط الملاحظتين الموجودتين في المنتصف. إن وسيط العينة أقل عرضة للتطرف مقارنة مع متوسط العينة. إنه يتجاهل الانحراف في مجتمع إحصائي من العينات أو في دالة كثافة

الاحتمال للمجتمع الإحصائي. بشكل عام، لتمثيل النزعة المركزية بوضوح لمجتمع من العينات أو المجتمع الإحصائي، فإننا نستخدم القاعدة التالية للإبهام:

١ – إذا كان المدرج التكراري للعينة أو دالة كثافة الاحتمال للمجتمع الإحصائي متماثلا، استخدم المتوسط كمقياس مركزي. بالنسبة لمثل هذه المجتمعات الإحصائية، فإن المتوسط والوسيط متساويين تقريباً، وتقدير المتوسط يستخدم كافة البيانات.

٢ – إذا كان المدرج التكراري للعينة أو دالة كثافة الاحتمال للمجتمع الإحصائي غير متماثل، فإن الوسيط هو المقياس الأكثر ملاءمة لمركز التوزيع.

المقياس الآخر للنزعة المركزية هو المنوال، الذي هو ببساطة الملاحظة الأكثر تكراراً في مجتمع العينات. في مثال الحجم المدّي المذكورة سابقاً، فإن العدد 1.3 هو قيمة العينة الأكثر تكراراً. لا يتم استخدام المنوال بنفس تكرار استخدام المتوسط أو الوسيط في تمثيل النزعة المركزية.

**(٣,٦,٢) مقاييس المتغير Measures of Variability**

تعتبر مقاييس النزعة المركزية وحدها غير كافية لتمثيل إحصائيات المجتمع الإحصائي أو العملية. من المألوف في الواقع وجود اختلاف في المجتمع الإحصائي الذي يجعل الأمور مثيرة للاهتمام ويؤدي إلى عدم اليقين في صنع القرار. إن الاختلاف من شخص إلى شخص، وخصوصاً في وظيفة فسيولوجية، هو ما يجعل إيجاد التشخيص والعلاج غير المضلل في كثير من الأحيان صعباً للغاية. إن ما يصلح لشخص واحد غالباً ما يفشل مع شخص آخر، وليس المتوسط أو الوسيط هو الذي يتم الحصول عليه من تلك الاختلافات من شخص إلى شخص، بل هو الاختلاف، الذي يتم عكسه في اختلاف نماذج الاحتمال التي تنطوي تحتها تلك المجتمعات الإحصائية المختلفة.

عند تلخيص اختلاف مجتمع أو عملية، فإننا نسأل عادة، "ما هو بعد موقع العينات (البيانات) عن المركز (متوسط العينة)؟". للإجابة على هذا السؤال، نستخدم

عادة التقديرات التالية التي تمثل انتشار بيانات النموذج: مدى الربيعى، تباين العينة، والانحراف المعياري للعينة.

إن مدى الربيعى هو الفرق بين الربيع الأول والثالث لبيانات العينة. وبالنسبة للبيانات التي تم أخذ عينات لها، فإن الوسيط معروف أيضاً بأنه الربيع الثاني، Q2. وبالأخذ في الاعتبار Q2، يمكننا إيجاد الربيع الأول، Q1، ببساطة من خلال أخذ قيمة الوسيط لتلك العينات التي تقع تحت الربيع الثاني. يمكننا إيجاد الربيع الثالث، Q3، بأخذ قيمة الوسيط لتلك العينات التي تقع فوق الربيع الثاني. وكتوضيح لذلك، لدينا العينات التالية:

1, 3, 3, 2, 5, 1, 1, 4, 3, 2,

إذا رتبنا هذه العينات ترتيباً تنازلياً،

5, 4, 3, 3, 3, 2, 2, 1, 1, 1,

إن قيمة الوسيط والربيع الثاني لهذه العينات هي 2.5. ويمكن إيجاد الربيع الأول، Q1، من خلال إيجاد وسط العينات التالية،

2.5, 2, 2, 1, 1, 1,

الذي يساوي 1.5. بالإضافة إلى ذلك، قد يتم إيجاد الربيع الثالث، Q3، بأخذ قيمة وسط العينات التالية:

5, 4, 3, 3, 3, 2.5,

الذي يساوي 3. وهكذا، فإن مدى الربيعي، Q3 – Q1 = 3 -1.5 = 1.5.

يتم تعريف تباين العينة، $s^2$، بأنه "متوسط المسافة للبيانات عن المتوسط"، وصيغة تقدير $s^2$ من مجتمع من العينات، $x_i$، هو:

$$s^2 = \frac{1}{n-1}\sum_{i-1}^{n}\left(x_i - \bar{x}\right)^2$$

إن الانحراف المعياري للعينة ، $s$، الذي غالباً ما يُشار إليه في وصف تغيّر البيانات هو :

$s=\sqrt{s^2}$ (نفس وحدات العينات الأصلية).

من المهم الإشارة إلى أن التوزيعات الطبيعية (المدرجات التكرارية المتماثلة)، ومتوسط العينة وانحراف العينة هي البارامترات الوحيدة اللازمة لوصف إحصائيات الظاهرة الأساسية. وهكذا، إذا كان ينبغي للمرء مقارنة مجموعتين أو أكثر موزعتين طبيعياً ، فإنه يحتاج فقط إلى اختبار تكافؤ المتوسطات والتباين لتلك المجتمعات الإحصائية.

# الفصل الرابع

## افتراض نموذج احتمال من بيانات العينة

## Assuming a Probability Model From the Sample Data

الآن بعد أن جمعنا البيانات، ورسمنا المدرج التكراري، وقدرنا المقاييس للنزعة المركزية والتغيّر، مثل المتوسط والوسيط والانحراف المعياري، فإننا جاهزين لافتراض نموذج احتمال للمجتمع أو العملية التي حصلنا منها على العينات. عند هذه النقطة، نقوم بافتراض تقريبي باستخدام مقادير بسيطة للمتوسط، والوسيط، والانحراف المعياري، والمدرج التكراري. ولكن من المهم ملاحظة أن هناك اختبارات أكثر صرامة، مثل اختبار $\chi^2$ للطبيعية [7] لتحديد ما إذا كان نموذج احتمال خاص مناسبًا للافتراض من مجتمع لبيانات العينة.

بمجرد افتراضنا لنموذج الاحتمال المناسب، يمكننا اختيار الاختبارات الإحصائية المناسبة التي ستتيح لنا اختبار الفروض واستخلاص النتائج بمستوى معين من الثقة. وسيحدد نموذج الاحتمال مستوى الثقة لدينا عند قبول أو رفض فرضية ما.

هناك سؤالان أساسيان نحاول معالجتهما عند افتراض نموذج احتمال لمجموعتنا الأساسية:

١ – ما هو مقدار ثقتنا بأن إحصائيات العينة تمثل المجتمع الإحصائي الكامل؟

٢- هل الاختلافات في الإحصائيات بين مجتمعين كبيرين، وناتجة عن عوامل أكثر منها عن الصدفة فقط؟

نحتاج، من أجل تأكيد أي مستوى "الثقة" عند القيام بالاستنتاج الإحصائي، إلى نموذج رياضي يصف احتمال إمكانية حدوث أي قيمة بيانات. وتُسمى هذه النماذج توزيعات الاحتمال.

هناك عدد من نماذج الاحتمال التي يتم افتراضها تكراراً لوصف العمليات البيولوجية. على سبيل المثال، عند وصف تغيّر معدل ضربات القلب، فإن احتمال ملاحظة فترة زمنية محددة بين ضربات القلب المتتالية يمكن وصفه عن طريق التوزيع الأسي [1, 8]. يوضح الشكل (٣.٦) في الفصل الثالث مدرجا تكرارياً (رسماً بيانياً) لعينات مأخوذة من توزيع أسي. لاحظ أن هذا التوزيع مائل بشدة إلى اليمين. وبالنسبة للفواصل R-R، فإن دالة الاحتمال هذه منطقية من الناحية الفسيولوجية لأن خلايا القلب الفردية لها فترة استعصاء تمنعها من التقلص في أقل من فترة زمنية دنيا. ومع ذلك، قد تحدث فترة زمنية طويلة جداً بين الضربات، مما يؤدي إلى ظهور بعض الفترات الزمنية الطويلة بشكل غير منتظم.

إن نموذج الاحتمال الأكثر افتراضاً لمعظم التطبيقات العلمية والهندسية هو التوزيع الطبيعي أو توزيع جاوس (normal or Gaussian distribution). وهذا التوزيع موضَّح بالخط الأسود المتصل في الشكل (٤.١) وغالباً ما يُشار إليه بمنحني الجرس لأنه يشبه الجرس الموسيقي.

إن المعادلة التي تعطي الاحتمال، $f(x)$ ، لملاحظة قيمة معينة لـ $x$ من المجتمع الإحصائي هي:

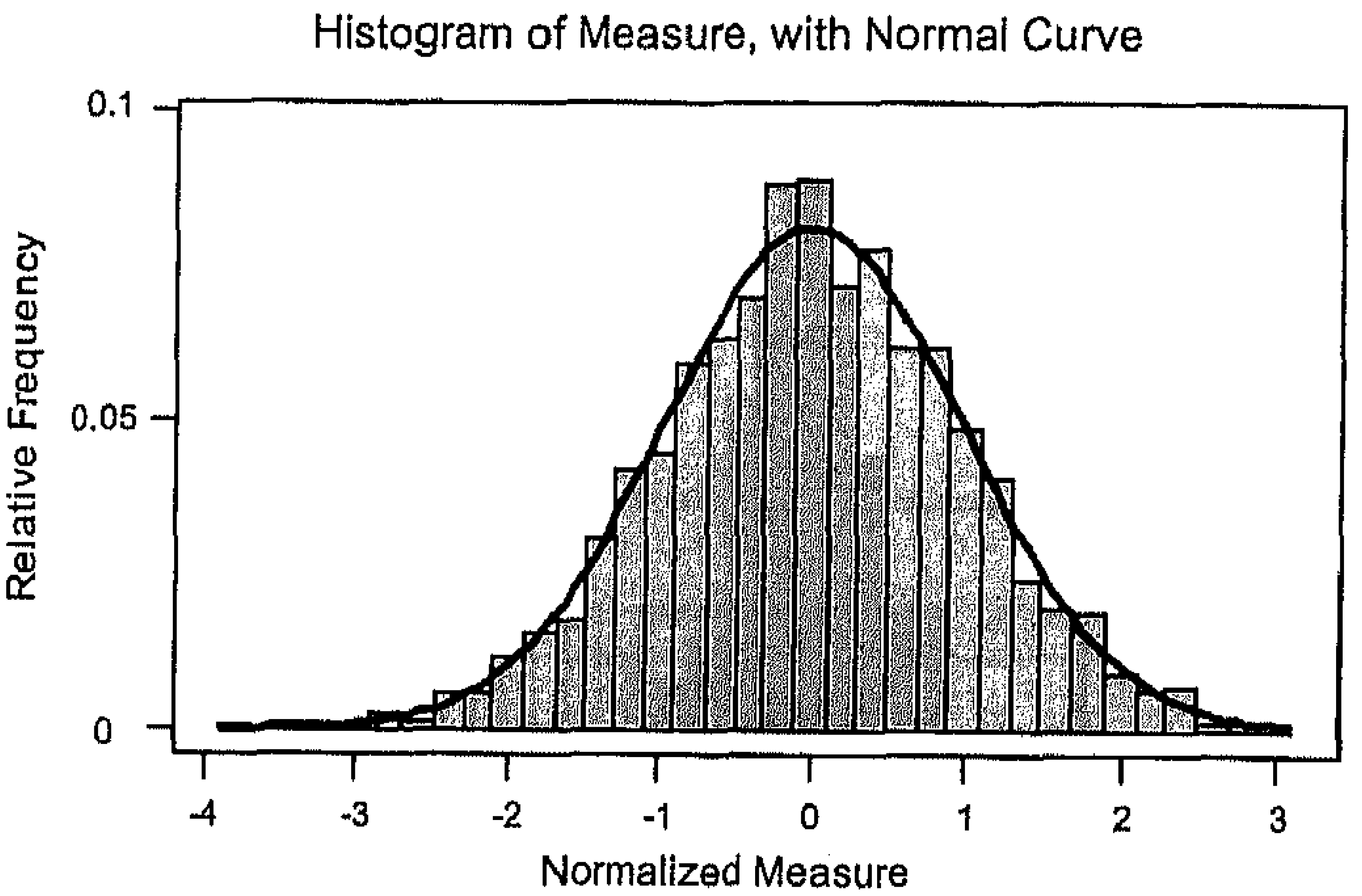
$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad \infty < x < \infty$$

حيث $\mu$ هو المتوسط الحقيقي للمجتمع أو العملية و $\sigma$ هو الانحراف المعياري لنفس المجتمع الإحصائي أو العملية. والرسم بياني لهذه المعادلة مُعطى مُوضَّحاً بالمنحنى المتصل الأملس في الشكل (٤,١). المساحة تحت المنحنى تساوي واحد.

لاحظ أن التوزيع الطبيعي هو:

١- منحني متناظر على شكل جرس موصوف تماماً بواسطة، متوسطه، $\mu$، وانحرافه المعياري، $\sigma$.

٢- من خلال تغيير $\mu$ و $\sigma$، نمد ونزيح التوزيع.

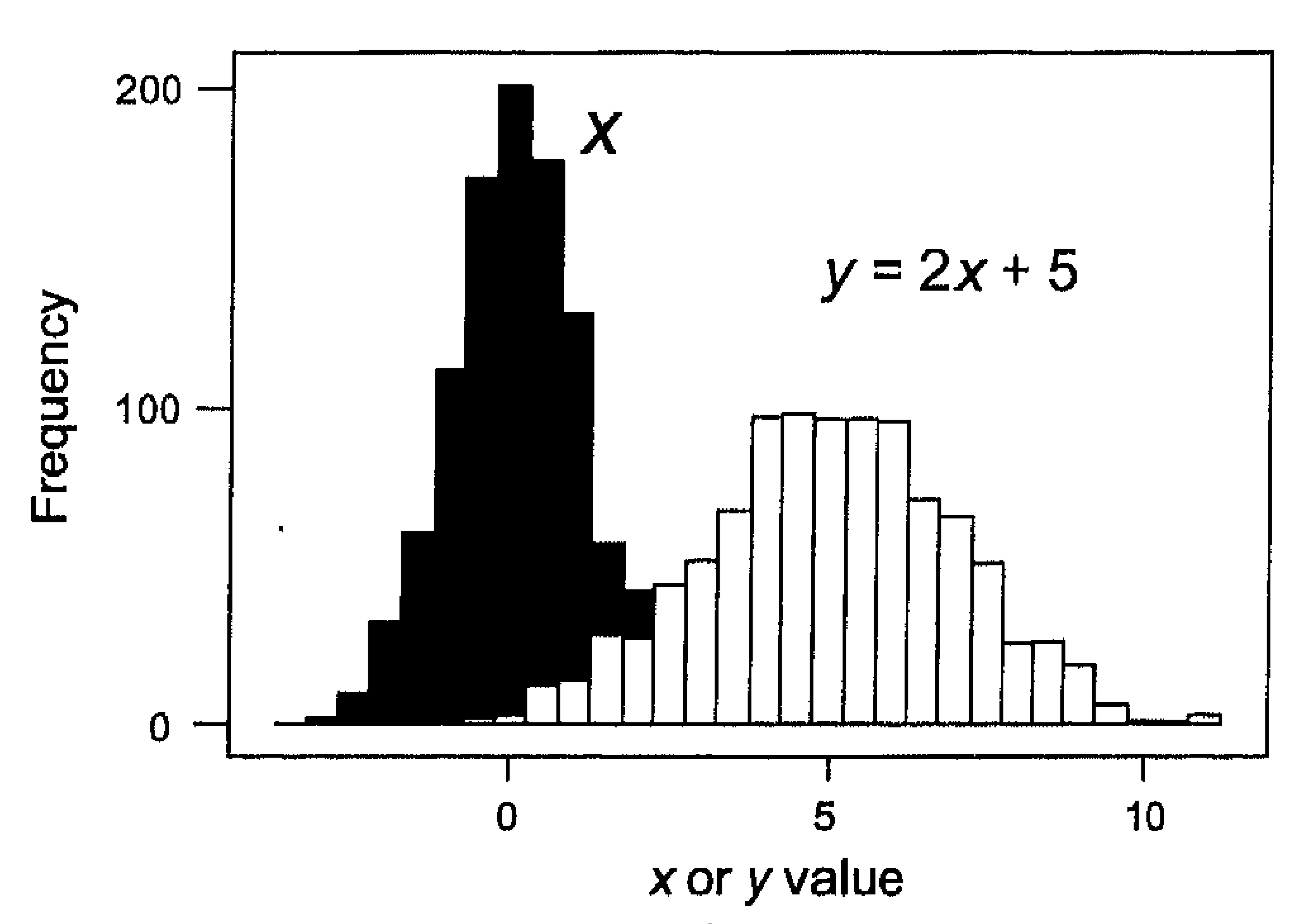



**الشكل (١ , ٤). يتم هنا توضيح المدرج التكراري لـ 1000 عينة مأخوذة من توزيع طبيعي. وتم وضع المنحنى الطبيعي المثالي الممثل لدالة توزيع الاحتمال الطبيعي على المدرج التكراري.**

يوضح الشكل (٤,١) أيضاً مدرجا تكرارياً يتم الحصول عليه عندما نختار عشوائياً 1000 عينة من المجتمع الإحصائي الموزعة بشكل طبيعي، ولها متوسط قيمته صفر تباين (variance) قيمته 1. ومن المهم الاعتراف أنه عند زيادة حجم العينة $n$، فإن المدرج التكراري يقترب من التوزيع الطبيعي المثالي المُبين بالخط المتصل الأملس. ولكن، عند أحجام عينة صغيرة، قد يبدو المدرج التكراري مختلفاً جداً عن المنحنى

الطبيعي. وهكذا، قد يكون من الصعب، من أحجام عينة صغيرة، تحديد ما إذا كان النموذج المُفترض مناسباً للمجتمع أو العملية، وأن أية اختبارات إحصائية نقوم بها قد لا تسمح لنا باختبار الفروض واستخلاص النتائج بأي مستوى من الثقة الحقيقية.

يمكننا القيام بعمليات خطية على المتغير العشوائي المُوزَّع بشكل طبيعي، $x$، لإنتاج متغير عشوائي آخر مُوزَّع بشكل طبيعي، $y$. تشمل هذه العمليات ضرب $x$ بثابت وإضافة ثابت (تعويض) إلى $x$. يوضح الشكل (٤،٢) المدرجات التكرارية لعينات مأخوذة من كل من المجموعتين $x$ و $y$. نلاحظ أن توزيع $y$ مُزاح (المتوسط يساوي الآن 5)، وأن التغيّر قد ازداد بالنسبة إلى $x$.




**الشكل (٢ ، ٤). مدرجات تكرارية مُبينة لعينات مأخوذة من مجتمعين $x$ و $y$، حيث $y$ هو مجرد دالة خطية لـ $x$. لاحظ أن المتوسط والتغيّر لـ $y$ يختلف عن $x$، ولكن كليهما توزيع طبيعي.**

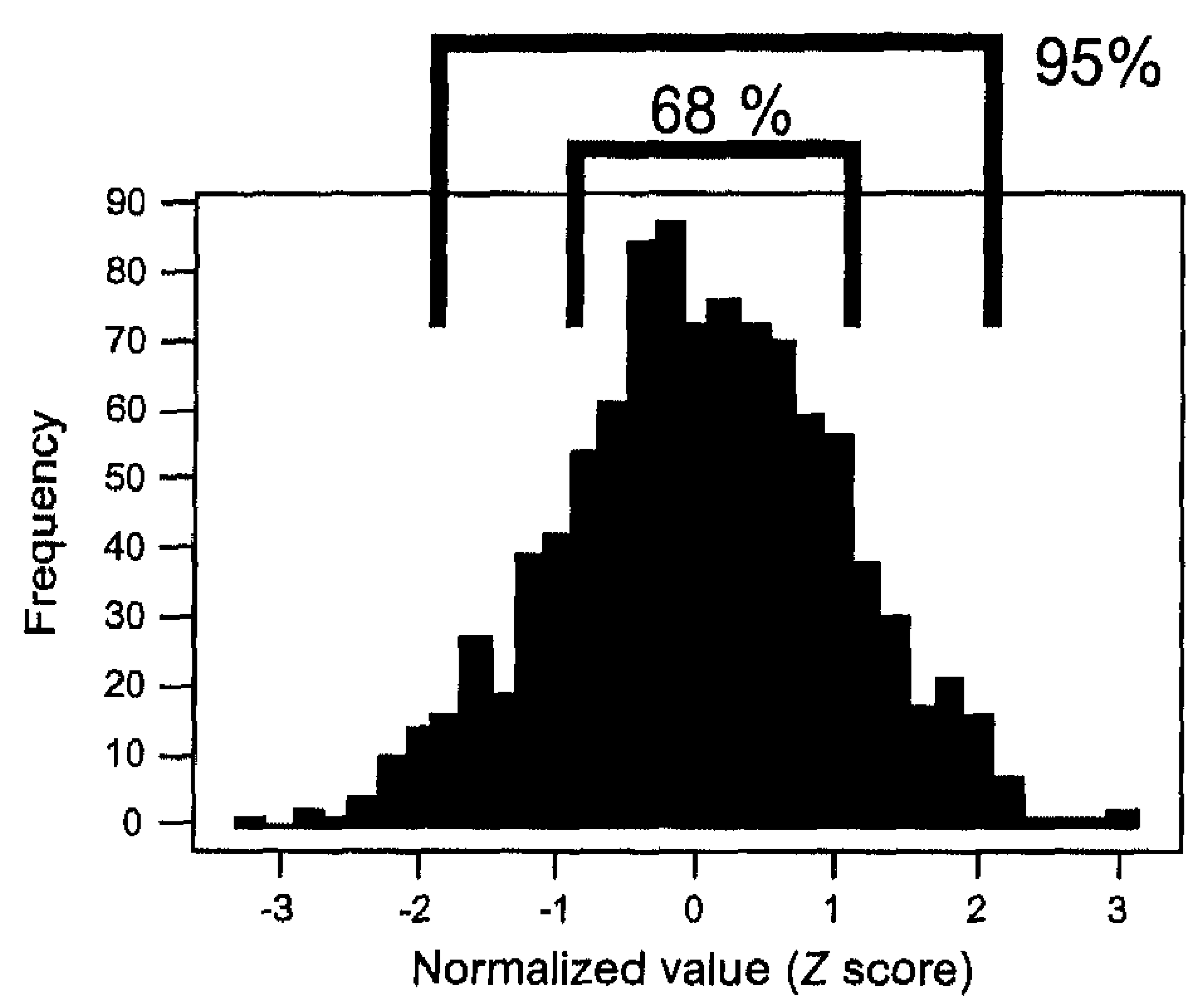
إن أحد الاختبارات التي قد نستخدمها لتحديد مدى جودة ملاءمة نموذج الاحتمال الطبيعي للبيانات المتوفرة لدينا هو عدد العينات التي تندرج ضمن الانحرافين المعياريين ±1 و ±2 للمتوسط. إذا كانت البيانات والمجتمع الإحصائي أو العملية مُنمذجة جيداً بواسطة التوزيع الطبيعي، فإن 68% من العينات ينبغي أن يقع ضمن الانحراف

المعياري ±1 عن المتوسط و 95% من العينات ينبغي أن تقع ضمن الانحرافات المعيارية ±2 عن المتوسط. هذه النسب مُوضَّحة في الشكل (٤,٣). ومن المهم تذكر هذه الأرقام القليلة، لأننا سوف نستخدم في كثير من الأحيان مجال الـ 95% عند استخلاص النتائج من التحليل الإحصائي.

هناك وسيلة أخرى لتحديد مدى جودة تمثيل بياناتنا التي تم أخذ عيناتها، $x$ ، لتوزيع طبيعي وهي تقدير معامل بيرسون للانحراف (عدم التماثل) (PCS) [5]. ويُعطى معامل الانحراف بواسطة المعادلة التالية:

$$PCS = \frac{3\left|\bar{x} - x_{median}\right|}{s}$$

إذا كان PCS > 0.5 ، فإننا نفترض أن العينات لم تكن مأخوذة من مجتمع مُوزَّع بشكل طبيعي.




**الشكل (٣ , ٤). مدرج تكراري لعينات مأخوذة مجتمع إحصائي مُوزَّع بشكل طبيعي. وبالنسبة للتوزيع الطبيعي، ينبغي لـ 68% من العينات أن تقع ضمن الانحراف المعياري ±1 عن المتوسط. (صفر في هذه الحالة) و 95% من العينات ينبغي أن تقع ضمن الانحرافات المعيارية ±2 (1.96 على وجه الدقة) عن المتوسط.**

عند جمع البيانات، فإنه عادة ما يتم جمعها بالعديد من الأنواع المختلفة من الوحدات الفيزيائية (فولت، مئوية، نيوتن، سنتمتر، غرام، ...إلخ). ومن أجل استخدام الجداول التي تم وضعها لنماذج احتمالات، فهناك حاجة إلى تنظيم البيانات بحيث يكون للبيانات المُنتظمة متوسط قيمته صفر، وانحراف معياري قيمته 1. ويُسمى مثل هذا التوزيع الطبيعي بالتوزيع الطبيعي المعياري وهو مُوضَّح في الشكل (٤,١).

التوزيع الطبيعي المعياري له توزيع متماثل على شكل جرس وفيه $\mu=0$ و $\sigma=1$.

لتحويل البيانات المُوزَّعة بشكل طبيعي إلى قيمة طبيعية معيارية، فإننا نستخدم الصيغ التالية:

$$z=(x-\mu)/\sigma \text{ أو } z=(x-\bar{x})/s,$$

وذلك اعتماداً على ما إذا كنا نعرف المتوسط الحقيقي، $\mu$، والانحراف المعياري، $\alpha$، أو ليس لدينا سوى تقديرات العينة، $\bar{x}$ أو $s$.

لأجل أية عينة فردية أو نقطة بيانات، $x_i$، من عينة ذات متوسط، $\bar{x}$، وانحراف معياري، $s$، يمكننا تحديد نقاطها $z$ من الصيغة التالية:

$$z_i=\frac{x_i-\bar{x}}{s}.$$

وبالنسبة لعينة فردية، فإن النقاط $z$ هي قيمة "منتظمة" أو "قياسية". يمكننا استخدام هذه القيمة مع معادلات دالة كثافة الاحتمال أو جداول الاحتمال القياسية [3] لتحديد احتمال ملاحظة مثل قيمة عينة من المجتمع.

يمكن التفكير أيضاً بالنقاط $z$ بأنها مقياس لمسافة العينة الفردية، $x_i$، عن متوسط العينة، $\bar{x}$، في وحدات الانحراف المعياري. على سبيل المثال، إذا كانت لنقطة عينة، $x_i$ $z$ نقطة مقدارها $z_i=2$، فهذا يعني أن نقطة البيانات، $x_i$، تبعد انحرافين معياريين عن متوسط العينة.

إننا نستخدم نقاط $z$ منتظمة بدلاً من البيانات الأصلية عند إجراء التحليل الإحصائي ؛ لأن جداول البيانات المنتظمة يتم تنظيمها بالفعل ومتوفرة في معظم كتب الإحصاء أو حزم البرامج الإحصائية. بالإضافة إلى ذلك، وباستخدام القيم المنتظمة، ليس هناك داعٍ للقلق بشأن المطال المطلق للبيانات أو الوحدات المستخدمة لقياس البيانات.

## (٤,١) التوزيع الطبيعي المعياري
## THE STANDARD NORMAL DISTRIBUTION

يوضح الجدول (٤,١) التوزيع الطبيعي المعياري.

يوفّر الجدول $z$ المرتبط بهذا الشكل مداخل الجدول الذي يعطي احتمال أن $z \le a$ ، الذي يساوي المساحة تحت المنحنى الطبيعي إلى يسار $z = a$. إذا كانت البيانات تأتي من توزيع طبيعي، فإن الجدول يخبرنا عن احتمال أو "فرصة" قيمة العينة أو النتائج التجريبية التي لها قيمة أقل من أو تساوي $a$.

وهكذا، يمكننا أخذ أي عينة وحساب نقاطها $z$ على النحو الموصوف سابقاً ومن ثم استخدم الجدول $z$ لإيجاد احتمال ملاحظة قيمة $z$ أقل من قيمة منتظمة محددة أو تساويها، $a$. على سبيل المثال، إن احتمال ملاحظة قيمة $z$ أقل من أو تساوي 1.96 هو 97.5%. وهكذا، فإن احتمال ملاحظة قيمة $z$ أكبر من 1.96 هو 2.5%. بالإضافة إلى ذلك، وبسبب التماثل في التوزيع، نعلم أن احتمال ملاحظة قيمة $z$ أقل من − 1.96 هو أيضاً 97.5%، واحتمال ملاحظة قيمة $z$ أقل من أو تساوي − 1.96 هو 2.5%. وأخيراً، فإن احتمال ملاحظة قيمة $z$ بين − 1.96 و 1.96 هو 95%. ينبغي للقارئ دراسة الجدول $z$ والرسم البياني المرتبط به للتوزيع $z$ للتحقق من أن الاحتمالات (أو المساحات الواقعة تحت دالة كثافة الاحتمال) المذكورة لاحقا صحيحة.

نحن بحاجة في كثير من الأحيان لتحديد احتمال بحيث تقع نتيجة تجريبية بين قيمتين أو تكون النتيجة أكبر من قيمة محددة $a$ أو أقل أو أكبر من قيمة محددة $b$. ولإيجاد هذه المساحات، يمكننا استخدام الصيغ الهامة التالية، حيث Pr هو الاحتمال:

$$\Pr(a \le z \le b) = \Pr(z \le b) - \Pr(z \le a)$$

المساحة بين $z = a$ و $z = b$

$$\Pr(z \le a) = 1 - \Pr(z < a)$$

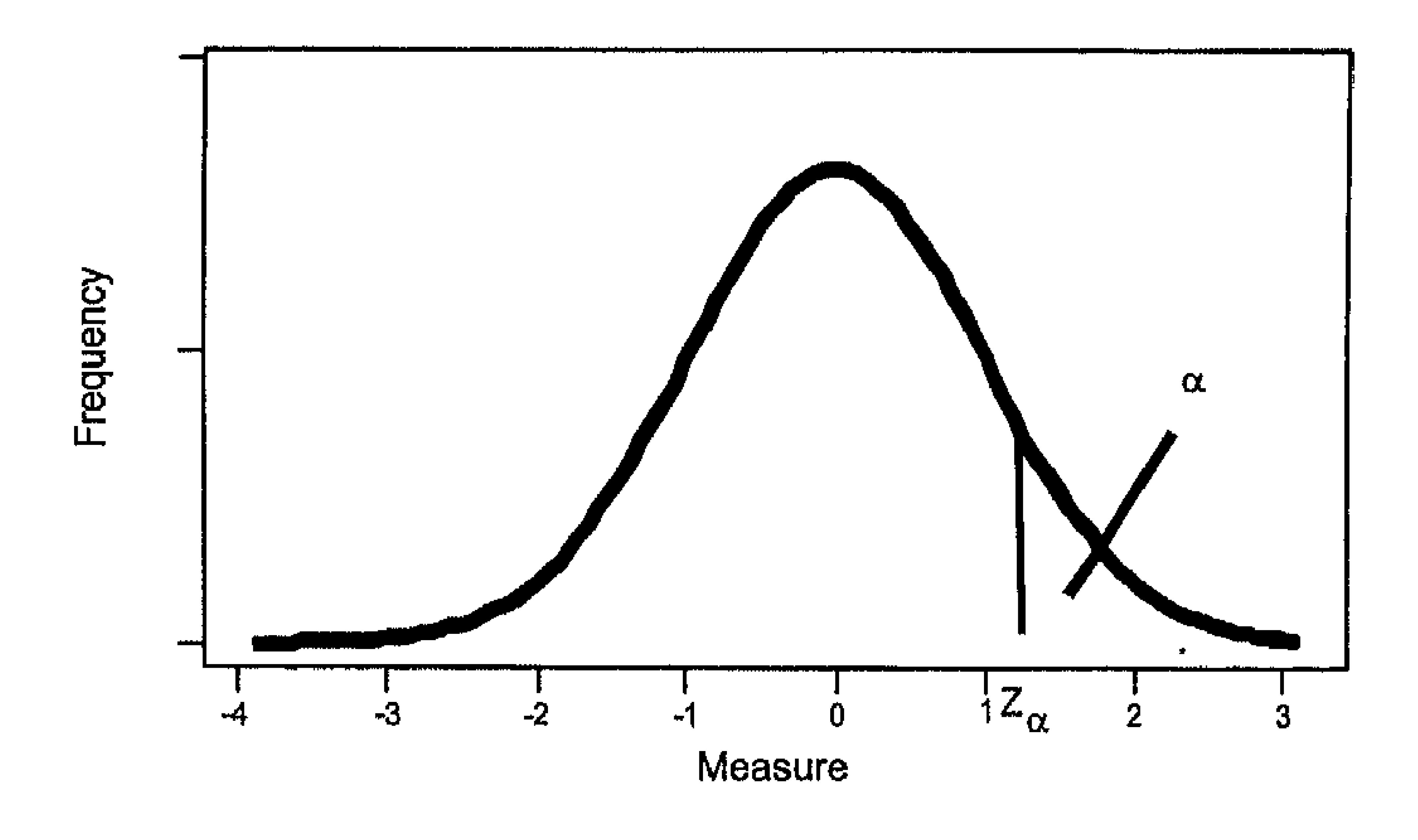
المساحة إلى يمين $z = a$

المساحة في الذيل الأيمن

وهكذا، من أجل أية ملاحظة أو قياس، $x$، من أي توزيع طبيعي:

$$\Pr(a \le x \le b) = \Pr\left(\frac{a-\mu}{\sigma} \le z \le \frac{b-\mu}{\sigma}\right),$$

حيث $\mu$ هو متوسط التوزيع الطبيعي و $\sigma$ هو الانحراف المعياري للتوزيع الطبيعي.



الجدول (١ , ٤). دالة توزيع z معياري: المساحات تحت دالة الكثافة الطبيعية القياسية.

| **0.09** | **0.08** | **0.07** | **0.06** | **0.05** | **0.04** | **0.03** | **0.02** | **0.01** | **0.00** | **Z** |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.5359 | 0.5319 | 0.5279 | 0.5239 | 0.5199 | 0.5160 | 0.5120 | 0.5080 | 0.5040 | 0.5000 | 0.0 |
| 0.5753 | 0.5714 | 0.5675 | 0.5636 | 0.5596 | 0.5557 | 0.5517 | 0.5478 | 0.5438 | 0.5398 | 0.1 |
| 0.6141 | 0.6103 | 0.6064 | 0.6026 | 0.5987 | 0.5948 | 0.5910 | 0.5871 | 0.5832 | 0.5793 | 0.2 |
| | | | | | | | | | | ... |
| 0.9633 | 0.9625 | 0.9616 | 0.9608 | 0.9599 | 0.9591 | 0.9582 | 0.9573 | 0.9564 | 0.9554 | 1.7 |
| 0.9706 | 0.9699 | 0.9693 | 0.9686 | 0.9678 | 0.9671 | 0.9664 | 0.9656 | 0.9649 | 0.9641 | 1.8 |
| 0.9767 | 0.9761 | 0.9756 | 0.9750 | 0.9744 | 0.9738 | 0.9732 | 0.9726 | 0.9719 | 0.9713 | 1.9 |
| 0.9817 | 0.9812 | 0.9808 | 0.9803 | 0.9798 | 0.9793 | 0.9788 | 0.9783 | 0.9778 | 0.9772 | 2.0 |
| | | | | | | | | | | ... |
| 0.9936 | 0.9934 | 0.9932 | 0.9931 | 0.9929 | 0.9927 | 0.9925 | 0.9922 | 0.9920 | 0.9918 | 2.4 |
| 0.9952 | 0.9951 | 0.9949 | 0.9948 | 0.9946 | 0.9945 | 0.9943 | 0.9941 | 0.9940 | 0.9938 | 2.5 |
| 0.9964 | 0.9963 | 0.9962 | 0.9961 | 0.9960 | 0.9959 | 0.9957 | 0.9956 | 0.9955 | 0.9953 | 2.6 |
| | | | | | | | | | | ... |
| 0.9990 | 0.9990 | 0.9989 | 0.9989 | 0.9989 | 0.9988 | 0.9988 | 0.9987 | 0.9987 | 0.9987 | 3.0 |

قيم z المُستخدمة بشكل شائع

| $Z_\alpha$ | المساحة في الذيل الأيمن، $\alpha$ |
|---|---|
| 1.282 | 0.1 |
| 1.645 | 0.05 |
| 1.96 | 0.025 |
| 2.326 | 0.010 |
| 2.576 | 0.005 |

وبعبارة أخرى، نحن بحاجة إلى معايرة أو إيجاد القيم $z$ لكل من البارامترات، $a$ و $b$، لإيجاد المساحة الواقعة تحت المنحنى الطبيعي المعياري (التوزيع $z$) التي تمثل التعبير على الجانب الأيسر من المعادلة المذكورة سابقاً.

**مثال (٤,١)**

إن متوسط امتصاص الدهون بالنسبة للذكور الذين تتراوح أعمارهم من 6 – 9 سنوات هو 28 غراما، بانحراف معياري مقداره 13.2 غرام. افترض أن الامتصاص مُوزع بشكل طبيعي. إن امتصاص ستيف (Steve) هو 42 غراما وامتصاص بن (Ben) هو 25 غراما.

١– ما هي نسبة المساحة بين امتصاص ستيف اليومي وامتصاص بن اليومي؟

٢– إذا أردنا عشوائياً اختيار ذكر يتراوح عمره بين 6 و 9 سنوات، ما هو احتمال أن يكون امتصاصه للدهون 50 غرام أو أكثر؟

**الحل**

$x$ = امتصاص الدهون

١– يمكن القول إن المشكلة على النحو التالي: ما هو $\Pr(25 \le x \le 42)$ ؟

بافتراض أن التوزيع طبيعي، فإننا نتحول إلى نقاط $z$:

ما هو $\Pr(((25-28)/13.2) < z < ((42-28/13.2)))$ ؟

$= \Pr(-0.227 \le z \le 1.06) = \Pr(z \le 1.06) - \Pr(z \le -0.227)$

(وذلك باستخدام صيغة $\Pr(a \le z \le b)$)

$\Pr(z \le 1.06) - [1 - \Pr(z \le 0.227)] = 0.8554 - [1 - 0.5910] =$.

$0.4464 \, or \, 44.6\%$ (من المساحة تحت المنحنى $z$)

٢– يمكن القول إن المشكلة على النحو التالي: "ما هو $\Pr(x > 50)$" ؟

بالمعايرة إلى النقطة $z$، ما هو $\Pr(z > (50-28/13.2))$ ؟

$= \Pr(z > 1.67)$

$= 1 - \Pr(z \le 1.67) = 1 - 0.9525 = 0.0475,$

$or \, 4.75\%$ من المساحة تحت المنحنى $z$.

**مثال (٤,٢)**

بفرض أن المواصفات على مجال إزاحة مُفرِّض (أداة تفريض) (indentor) طرف هو 0.001±0.5 ميللي متر. إذا كانت هذه الإزاحات مُوزَّعة بشكل طبيعي، ولها متوسط = 0.47 وانحراف معياري = 0.002، ما هي النسبة المئوية للمُفرِّضات الواقعة ضمن هذه المواصفات ؟

**الحل**

$x$ = الإزاحة

يمكن القول إن المشكلة على النحو التالي : "ما هو $\Pr(0.499 \le x \le 0.501)$"؟

باستخدام نقاط $z$،

$$\Pr(0.499 \le x \le 0.501) = \Pr((0.499 - 0.47)/0.002 \le z \le (0.501 - 0.47/0.002))$$

$$= \Pr(14.5 \le x \le 15.5) = \Pr(z \le 15.5) - \Pr(z \le 14.5) = 1 - 1 = 0$$

ومن المفيد ملاحظة أنه إذا كان توزيع المجتمع الإحصائي وبيانات العينة المرتبطة بها ليس طبيعياً (غير متماثل)، فإنه يمكن في كثير من الأحيان استخدام التحويلات لجعل البيانات طبيعية، ويمكن بعد ذلك استخدام الإحصائيات التي يغطيها هذا الكتاب لإجراء التحليل الإحصائي على البيانات التي تم تحويلها. وتشمل هذه التحويلات على البيانات الخام اللوغاريتم، والجذر التربيعي، والمقلوب (المعكوس).

## (٤,٢) التوزيع الطبيعي ومتوسط العينة
## THE NORMAL DISTRIBUTION AND SAMPLE MEAN

إن جميع التحليل الإحصائي يتبع نفس الإجراء العام :

١ – افتراض توزيع أساسي للبيانات والبارامترات المرتبطة به (على سبيل المثال، متوسط العينة).

٢- تدريج البيانات أو البارامتر بالنسبة إلى توزيع "معياري".

٣- تقدير فترات ثقة باستخدام جدول معياري للتوزيع المُفترض. (والسؤال الذي نطرحه هو: "ما هو احتمال ملاحظة النتيجة التجريبية التجريبية عن طريق الصدفة وحدها؟")

٤- إجراء اختبار الفرضية (على سبيل المثال، الاختبار $t$ للطالب).

بدأنا هذا الكتاب وركزنا في معظمه على التوزيع الطبيعي أو نموذج الاحتمال بسبب انتشاره في المجال الطبي الحيوي وإلى حد ما يُسمى نظرية النزعة المركزية. إن أحد أكثر الاختبارات الإحصائية الأساسية التي نقوم بها هو المقارنة بين المتوسطات لمجتمعين أو أكثر. إن متوسط العينة هو في حد ذاته تقدير تم إيجاده من عدد محدود من العينات. وهكذا، فإن متوسط العينة، $\bar{x}$، هو في حد ذاته متغير عشوائي تتم نمذجته بتوزيع طبيعي [4].

هل هذا النموذج لـ $\bar{x}$ منطقي؟ الجواب نعم، لعينات كبيرة، بسبب نظرية النزعة المركزية، والتي تنص على [4, 10]:

> إذا كانت نقاط أو عينات البيانات الفردية (كل عينة هي متغير عشوائي)، $x_i$، تأتي من أي توزيع احتمال عشوائي، فإن مجموع (وبالتالي، متوسط) نقاط البيانات تلك يكون مُوزَّعًا بشكل طبيعي عندما يصبح حجم العينة، $n$، كبيراً.

وهكذا، حتى إذا كانت كل عينة، مثل رمية النرد (الزهر)، تأتي من توزيع غير طبيعي (nonnormal) (على سبيل المثال، التوزيع المنتظم، مثل رمية النرد)، فإن مجموع تلك العينات الفردية (مثل المجموع الذي نستخدمه لتقدير متوسط العينة، $\bar{x}$) سيكون له توزيع طبيعي. ويمكن للمرء بسهولة افتراض أن العديد من العمليات البيولوجية أو الفسيولوجية التي نقيسها هو مجموع عدد من العمليات العشوائية مع توزيعات احتمال مختلفة، وبذلك، فإن الافتراض بأن العينات تأتي من توزيع طبيعي هو افتراض معقول.

## (٤,٣) فترة الثقة لمتوسط العينة
## CONFIDENCE INTERVAL FOR THE SAMPLE MEAN

كل إحصاء عينة في حد ذاته متغير عشوائي بنوع محدد لتوزيع الاحتمال. وهكذا، عندما نستخدم العينات لتقدير الإحصائيات الحقيقية للمجتمع الإحصائي (عادة ما تكون في الواقع غير معروفة ولا يمكن الحصول عليها)، فإننا نريد أن يكون لدينا مستوى معين من الثقة بأن تقديراتنا للعينة قريبة من إحصائيات المجتمع الإحصائي الحقيقية أو تكون ممثلة للمجتمع أو العملية.

عند تقدير فترة الثقة لمتوسط العينة، نطرح السؤال التالي: "ما هو مقدار تطابق متوسط العينة (مُقدَّراً من عدد محدود من العينات) مع المتوسط الحقيقي للمجتمع؟"

لتحديد مستوى الثقة للتقديرات الإحصائية أو الاستنتاجات الإحصائية، هناك حاجة أولاً إلى افتراض توزيع أو نموذج احتمال للعينات والمجتمع الإحصائي ومن ثم حاجة إلى تقدير فترة الثقة باستخدام التوزيع المُفترض.

إن فترة الثقة الأبسط الذي يمكننا أن نبدأ به فيما يتعلق بإحصائياتنا الوصفية هي فترة الثقة لمتوسط العينة، $\bar{x}$. سؤالنا هو: ما هو مدى تطابق متوسط العينة مع المتوسط الحقيقي للمجتمع أو العملية، $\mu$ ؟

قبل أن نتمكن من الإجابة على هذا، نحن بحاجة إلى افتراض نموذج احتمال أساسي لمتوسط العينة، $\bar{x}$ ، والمتوسط الحقيقي، $\mu$. وكما ذُكر سابقاً، يمكن إظهار أن متوسط العينة، $\bar{x}$ ، لحجم العينات الكبير مُنمذجاً جيداً من خلال التوزيع الطبيعي أو نموذج الاحتمال. وهكذا، سوف نستخدم هذا النموذج عند تقدير فترة الثقة لمتوسط العينة، $\bar{x}$ .

وهكذا يتم تقدير $\bar{x}$ من العينة. ثم نسأل، ما هو مدى تطابق $\bar{x}$ (متوسط العينة) مع المتوسط الحقيقي للمجتمع، $\mu$ ؟

يمكن إظهار أنه إذا أخذنا عدة مجموعات لـ $n$ عينة وقدَّرنا $\bar{x}$ لكل مجتمع ،

١- المتوسط أو متوسط العينة لـ $\bar{x} = \mu$ ، و

٢- الانحراف المعياري لـ $\bar{x} = s/\sqrt{n}$ .

وهكذا، عندما يصبح حجم العينة ، $n$، كبيراً ، فإن توزيع $\bar{x}$ يقترب من التوزيع الطبيعي.

عندما تكون $n$ كبيرة ، فإن $\bar{x}$ يتبع التوزيع الطبيعي ، وقد يتم استخدام النقطة $z$ لـ $\bar{x}$ لتقدير ما يلي :

$$\Pr(a \le \bar{x} \le b) = \Pr\left(\frac{a-\mu}{\sigma/\sqrt{n}} \le z \le \frac{b-\mu}{\sigma/\sqrt{n}}\right).$$

يفترض هذا التعبير أن $n$ كبيرة وأننا نعرف قيمة $\sigma$ .

ننظر الآن إلى الحالة التي قد تكون فيها $n$ كبيرة ، ولكننا لا نعرف قيمة. في مثل هذه الحالات ، نستبدل $\sigma$ بـ $s$ للحصول على التعبير التالي :

$$\Pr(a \le \bar{x} \le b) = \Pr\left(\frac{a-\mu}{s/\sqrt{n}} \le z \le \frac{b-\mu}{s/\sqrt{n}}\right),$$

حيث يُسمى $s/\sqrt{n}$ الخطأ المعياري للعينة ويمثل الانحراف المعياري لـ $\bar{x}$ .

لنفترض الآن $n$ كبيرة ، ونريد تقدير فترة الثقة 95% لـ $\bar{x}$. نقوم أولاً بتدريج متوسط العينة ، $\bar{x}$ ، بالنسبة لقيمة $z$ (لأن نظرية النزعة المركزية تنص على أن $\bar{x}$ مُوزَّعة بشكل طبيعي)

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

نحن نتذكر أن 95% من قيم $z$ تقع بين $\pm 1.96$ (حوالي $2\sigma$) من المتوسط ، وبالنسبة للتوزيع $z$،

$$\Pr(-1.96 \le z \le 1.96) = 0.95$$

وبتعويض $z$ المساوية

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

نحصل على

$$0.95 = \Pr\left(-1.96 \leq \frac{\bar{x} - \mu}{s/\sqrt{n}} \leq 1.96\right)$$

إذا استخدمنا الرموز التالية من حيث الخطأ المعياري للعينة :

$$SE(\bar{x}) = \frac{s}{\sqrt{n}}$$

بإعادة ترتيب حدود الصيغة سابقاً، نلاحظ أن احتمال وقوع $\mu$ بين $\pm$ 1.96 (أو 2) انحراف معياري لـ $\bar{x}$ هو 95%:

$$0.95 = \Pr\left(\bar{x} - 1.96SE(\bar{x}) \leq \mu \leq \bar{x} + 1.96SE(\bar{x})\right)$$

لاحظ أنه يُشار إلى 1.96 بـ $z_{\alpha/2}$ . هذه القيمة لـ $z$ هي قيمة $z$ التي من أجلها تكون المساحة في الذيل الأيمن للتوزيع الطبيعي مساوية $\alpha/2$ . إذا أردنا تقدير فترة ثقة 99%، فإننا نستبدل $z_{0.01/2}$ ، التي هي 2.576، في الموضع 1.96 أعلاه.

وهكذا، بالنسبة لـ $n$ كبيرة وأي مستوى وثوقية، $1-\alpha$ ، فإن فترة الثقة $1-\alpha$ للمتوسط الحقيقي للمجتمع الإحصائي، $\mu$ ، يُعطى بواسطة المعادلة التالية :

$$\mu = \bar{x} \pm z_{\alpha/2} SE(\bar{x})$$

وهذا يعني أن هناك $(1-\alpha)$ في المئة احتمال أن المتوسط الحقيقي يقع ضمن المجال أعلاه المتمركز حول $\bar{x}$ .

**مثال (٤,٣)**

تقدير فترات الثقة

المسألة : مُعطى مجتمع بيانات لها $\bar{x} = 505$ و $s$=100 . إذا كان عدد العينات 1000، ما هي فترة الثقة 95% بالنسبة لمتوسط المجتمع، $\mu$ ؟

**الحل**

إذا افترضنا حجم عينة كبير، يمكننا استخدام توزيع $z$ لتقدير فترة الثقة لمتوسط العينة باستخدام المعادلة التالية :

$$\mu = \bar{x} \pm z_{\alpha/2} SE(\bar{x})$$

وبالتعويض بالقيم التالية :

$$\bar{x} = 505;$$
$$SE(\bar{x}) = s/\sqrt{n} = 100/\sqrt{1000}$$

وبالنسبة فترة ثقة مقدارها 95%، فإن $\alpha = 0.05$.

باستخدام الجدول $z$ لتحديد موقع $z(\alpha/2)$ ، نجد أن قيمة $z$ التي تعطي مساحة قيمتها 0.025 في الذيل الأيمن هي 1.96

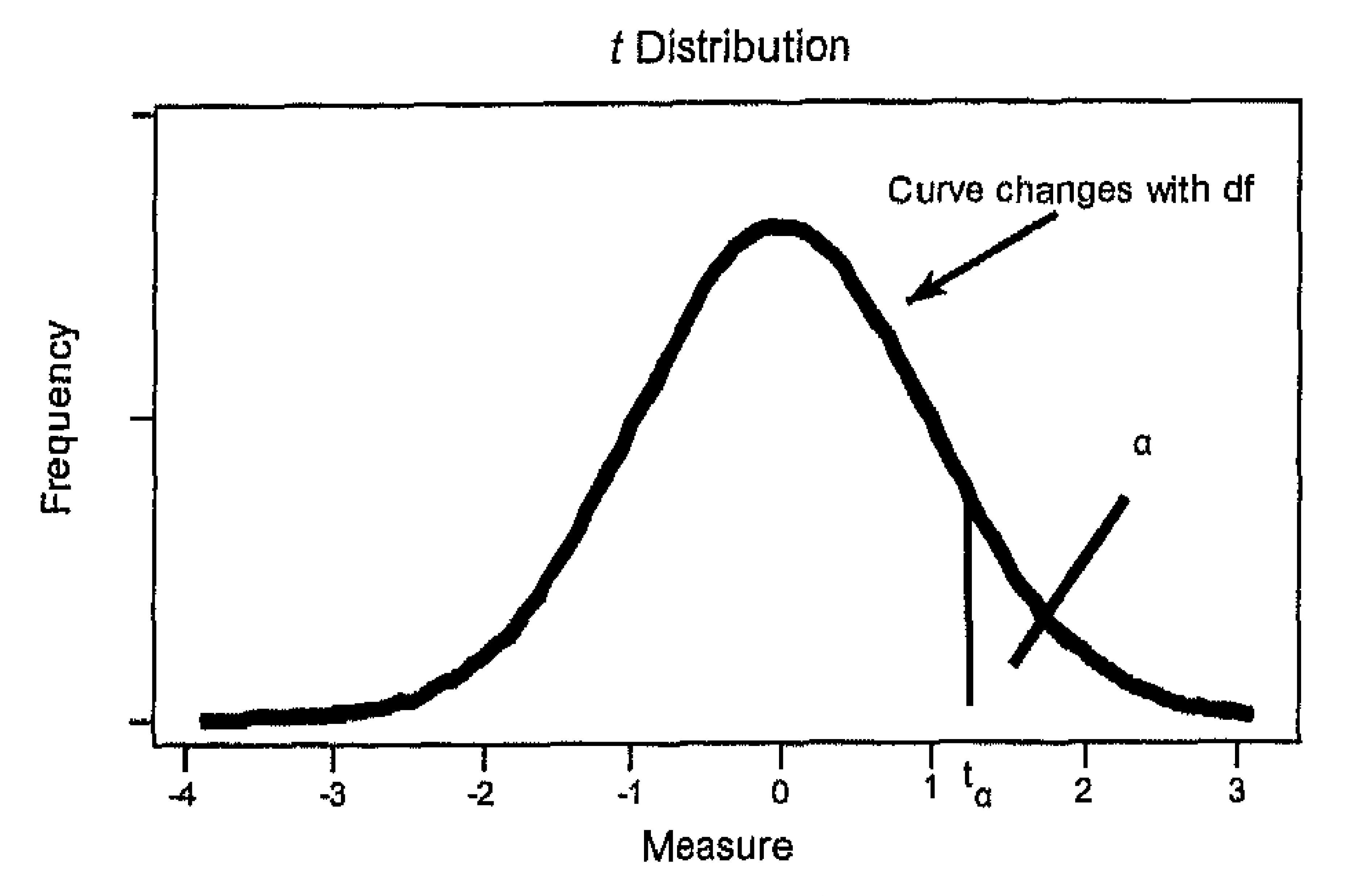
وبتعويض $\bar{x}$ ، و$SE(\bar{x})$ ، و$z(\alpha/2)$ في تقدير فترة الثقة أعلاه، نجد أن مدى الثقة 95% عندما $\mu = [498.80, 511.20]$.

لاحظ أنه إذا أردنا تقدير فترة ثقة مقدارها 99%، فإننا ببساطة نستخدم قيمة مختلفة لـ $z$، $z(0.01/2)$ في المعادلة نفسها. إن قيمة $z$ المرتبطة بمساحة مقدارها 0.005 في الذيل الأيمن تساوي 2.576. إذا استخدمنا هذه القيمة لـ $z$، فإننا نقدر فترة الثقة لـ $\mu$ بـ [496.86, 515.14]. ونلاحظ أن فترة الثقة قد اتسعت عند زيادة مستوى الثقة.

## (٤,٤) التوزيع $t$

## THE $t$ DISTRIBUTION

بالنسبة للعينات الصغيرة، لم يعد $\bar{x}$ مُوزَّعاً بشكل طبيعي. ولذلك، فإننا نستخدم التوزيع $t$ للطالب لتقدير الإحصاءات الحقيقية للمجتمع. يشبه التوزيع $t$، كما هو موضح في الجدول (٤,٢) التوزيع $z$ ولكن مع تناقص أبطأ عند الذيول ومنطقة وسطى أكثر تسطحاً.




مدخل الجدول = $t(\alpha;df)$، حيث $\alpha$ هي المساحة في الذيل إلى يمين $t(\alpha,df)$ و $df$ هي درجات الحرية.

الجدول (٢ , ٤). دالة النقط المئوية للتوزيع $t$ للطالب.

| $\alpha$ = المساحة إلى يمين $t(\alpha,df)$ | | | | | *df* |
|---|---|---|---|---|---|
| **0.005** | **0.01** | **0.025** | **0.05** | **0.10** | |
| 63.657 | 31.821 | 12.706 | 6.314 | 3.078 | 1 |
| 9.925 | 6.965 | 4.303 | 2.920 | 1.886 | 2 |
| 5.841 | 4.541 | 3.182 | 2.353 | 1.638 | 3 |
| 4.604 | 3.747 | 2.776 | 2.132 | 1.533 | 4 |
| 4.032 | 3.365 | 2.571 | 2.015 | 1.476 | 5 |
| | | | | | ... |
| 3.169 | 2.764 | 2.228 | 1.812 | 1.372 | 10 |
| 3.106 | 2.718 | 2.201 | 1.796 | 1.363 | 11 |
| 3.055 | 2.681 | 2.179 | 1.782 | 1.356 | 12 |
| 3.012 | 2.650 | 2.160 | 1.771 | 1.350 | 13 |
| 2.977 | 2.624 | 2.145 | 1.761 | 1.345 | 14 |
| 2.947 | 2.602 | 2.131 | 1.753 | 1.341 | 15 |
| | | | | | ... |
| 2.750 | 2.457 | 2.042 | 1.697 | 1.310 | 30 |
| 2.704 | 2.423 | 2.021 | 1.684 | 1.303 | 40 |
| 2.660 | 2.390 | 2.000 | 1.671 | 1.296 | 60 |
| 2.617 | 2.358 | 1.960 | 1.658 | 1.289 | 120 |
| 2.576 | 2.326 | 1.960 | 1.645 | 1.282 | ∞ |
| | | | | | |
| 2.576 | 2.326 | 1.960 | 1.645 | 1.282 | $z_\alpha$ (عينة كبيرة) |

نستخدم التوزيع $t$، مع ذيول متناقصة ببطء، لأنه مع العينات القليلة، لدينا يقين أقل حول التوزيع الأساسي (نموذج الاحتمال).

تُعطى الآن القيمة التي تم معايرتها من أجل $\bar{x}$ بواسطة العلاقة التالية:

$$\frac{\bar{x}-\mu}{s/\sqrt{n}}$$

المعروفة أن لها توزيعاً $t$ بدلاً من التوزيع $z$ الذي ناقشناه حتى الآن. وكان أول من اكتشف التوزيع $t$ W.S. Gosset [4, 11]، وهو كيميائي عمل في مصنع للجعة. قرر Gosset نشر التوزيع $t$ الخاص به تحت الاسم المستعار "الطالب". لذلك، فإننا غالباً ما نشير إلى هذا التوزيع بالتوزيع $t$ للطالب.

يعد التوزيع $t$ متماثلاً مثل التوزيع $z$ وله بشكل عام شكل الجرس. ولكن كمية انتشار التوزيع إلى الذيول، أو عرض الجرس، تعتمد على حجم العينة، $n$. وخلافاً للتوزيع $z$، الذي يفترض حجم عينة لا نهائية، فإن التوزيع $t$ يغير الشكل مع حجم العينة. والنتيجة هي أن فترات الثقة التي يتم تقديرها بالقيم $t$ هي أكثر "انتشاراً" منها للتوزيع $z$، وخاصة بالنسبة لأحجام عينة صغيرة، لأنه مع مثل هذه الأحجام للعينات، يتم لومنا لعدم وجود عينات كافية لتمثيل مدى التغير في المجتمع الإحصائي أو العملية. وهكذا، عند تقدير فترات الثقة لمتوسط العينة، $\bar{x}$ ، فإنه لا يوجد لدينا ثقة كبيرة في تقديراتنا. وبذلك، فإن الفترة تتسع لتعكس هذا اليقين المنخفض بأحجام العينة الأصغر. نقدر في المقطع التالي فترات الثقة لنفس المثال المُعطى مسبقاً، ولكن باستخدام التوزيع $t$ بدلاً من التوزيع $z$.

## (٤,٥) فترات الثقة باستخدام التوزيع $t$

## CONFIDANCE INTERVAL USING $t$ DISTRIBUTION

هناك جداول $t$، مثل الجداول $z$، تكون فيها القيم $t$ المرتبطة بمساحات مختلفة تحت منحنى الاحتمال محسوبة مسبقاً، ويمكن استخدامها للتحليل الإحصائي دون الحاجة إلى

إعادة حساب القيم $t$. إن الفرق بين الجدول $z$ والجدول $t$ هو أن القيم $t$ الآن دالة في الحجم العينات أو درجات الحرية. يعطي الجدول (٤،٢) بضعة أسطر من الجدول $t$ من [3].

لاستخدام الجدول ينظر المرء ببساطة إلى نقطة تقاطع درجات الحرية، $df$، (تتعلق بحجم العينة)، وقيمة $\alpha$ التي يرغب فيها المرء في الذيل الأيمن. توفر نقطة التقاطع القيمة $t$ التي من أجلها تكون المساحة تحت المنحنى $t$ في الذيل الأيمن مساوية $\alpha$. وبعبارة أخرى، فإن احتمال أن يكون $t$ أقل من إدخال محدد أو يساويه في الجدول هو $1-\alpha$. وبالنسبة لحجم عينة محدد، $n$، فإن درجات الحرية هي $df=n-1$.

نستبدل الآن ببساطة $t$ بـ $z$ لإيجاد فترات الثقة. لذا، فإن فترات الثقة لمتوسط العينة، $\overline{x}$، باستخدام التوزيع $t$ يصبح الآن:

$$\mu=\overline{x}\pm t\left(\frac{\alpha}{2};n-1\right)SE(\overline{x})$$

حيث $\mu$ هو المتوسط الحقيقي للمجتمع أو العملية التي نأخذ منها العينات، و $SE(\overline{x})$ هو الخطأ المعياري للمجتمع أو العملية، و $t$ هي قيمة $t$ التي يوجد من أجلها مساحة قيمتها $\alpha/2$ في الذيل الأيمن، و $n$ هو حجم العينة.

**المثال (٤،٤)**

فترة الثقة باستخدام التوزيع $t$.

**المسألة**

نأخذ في الاعتبار المثال نفسه المُستخدم سابقاً لتقدير فترات الثقة باستخدام القيم $z$. في هذه الحالة، يكون حجم العينة صغيرًا $n=20$، لذلك نستخدم الآن التوزيع $t$.

**الحل**

تقديرنا الآن فترة الثقة بالنسبة لـ $\mu=\overline{x}\pm t\left(\frac{\alpha}{2};n-1\right)SE(\overline{x})$

مرة أخرى، نستبدل بالقيم التالية:

$$\bar{x} = 505;$$
$$SE(\bar{x}) = s/\sqrt{n} = 100/\sqrt{20}$$

وبالنسبة لفترة ثقة مقدارها 95%، فإن $\alpha = 0.05$.

باستخدام الجدول $t$ لتحديد موقع $t(0.05/2, 20-1)$، نجد أنه من أجل 19 درجة حرية ($19\ df$)، أن قيمة $t$ التي تعطي مساحة قيمتها 0.025 في الذيل الأيمن هي 2.093.

وبتعويض $\bar{x}$، و $SE(\bar{x})$، و $t(\alpha, n-1)$ في تقدير فترة الثقة سابقاً، نجد أن فترة الثقة 95% عندما $\mu = [458.20, 551.80]$. ونلاحظ أن فترة الثقة هذه مُتسعة مقارنة مع الفترة الذي تم تقديرها سابقاً باستخدام القيم $z$. وهذا مُتوقع لأن التوزيع $t$ أوسع من التوزيع $z$ في أحجام العينة الصغيرة، مما يعكس حقيقة أن لدينا ثقة أقل في تقديرنا لـ $\bar{x}$ وبالتالي، $\mu$ عندما يكون حجم العينة صغيرًا.

ويمكن تقدير فترات الثقة لمعظم الإحصائيات الوصفية، مثل متوسط العينة، تباين العينة، وحتى الخط لأفضل ملاءمة مُحددة من خلال خط الانحدار [3]. وكما أشير سابقاً، يعكس فترة الثقة التغيّر في تقدير البارامتر وهو متأثر بحجم العينة، وتغيّر المجتمع، ومستوى الثقة الذي ننشده. كلما كانت الثقة المطلوبة أكبر، كانت الفترة أوسع. وبالمثل كلما كان التغير في العينات أكبر، كانت فترة الثقة أوسع.

**المثال (٥,٤)**

مراجعة مفاهيم الاحتمال.

**المسألة**

تم قياس أزمنة التجميع لعينة مؤلفة من 15 مضخة حقن غلوكوز. كان متوسط الزمن لتجميع مضخة حقن غلوكوز 15.8 دقيقة، مع انحراف معياري مقداره 2.4 دقيقة. بافتراض توزيع متماثل نسبي لأزمنة التجميع.

١ – ما النسبة المئوية لمضخات الحقن التي تتطلب أكثر من 17 ثانية للتجميع ؟

٢– ما فترة الثقة %99 للمتوسط الحقيقي لزمن التجميع $(\mu)$ ؟

٣– ما فترة الثقة %99 لمتوسط زمن التجميع إذا كان حجم العينة 2500؟

**الحل**

١ – $x$ = زمن التجميع.

ما هو $\Pr(x > 17)$ ؟

$\Pr(x > 17) = \Pr(z > (17 - 15.8)/2.4) = 1 - \Pr(z \leq 0.5) = 1 - 0.6915 = 0.3085$ ،

أو %38.50 من مضخات الحقن

٢– $= \bar{x} \pm t(\alpha/2; n-1) SE(x)$

$$= 15.8 \pm t((0.01)/2; 15-1)(2.4\sqrt{15})$$
$$= 15.8 \pm 2.977(0.6196)$$
$$= [13.96, 17.64]$$

لأن حجم العينة 2500 كبير الآن ، فإننا نستخدم قيمة $z$ لتقدير فترة الثقة ، $\mu = \bar{x} \pm z(\alpha/2) SE(x)$.

٣– $= \bar{x} \pm t(\alpha/2) SE(x)$

$$= 15.8 \pm t((0.01)/2)(2.4\sqrt{2500})$$

$$= 15.8 \pm 2.576(0.048)$$

$$= [15.68, 15.92]$$

# الفصل الخامس

## الاستنتاج الإحصائي
## Statistical Inference

الآن وبعد أن جمعنا البيانات وقدّرنا بعض الإحصائيات الوصفية الأساسية وافترضنا نموذج احتمال للمجتمع الإحصائي أو العملية فإننا جاهزون لإجراء بعض التحليلات الإحصائية التي من شأنها أن تسمح لنا بمقارنة المجتمعات الإحصائية واختبار الفروض. سنناقش في هذا الكتاب فقط التحليل الإحصائي الذي يكون فيه التوزيع الطبيعي نموذج احتمال جيد للمجتمع (المجتمعات الإحصائية). إذا كانت البيانات أو العينات تشير إلى أن المجتمع الإحصائي أو العملية ليست مُنمذجة بشكل جيد بواسطة التوزيع الطبيعي فإننا بحاجة عندئذ إلى اللجوء إلى أنواع أخرى من التحليل الإحصائي مثل التقنيات غير البارامترية (nonparametric) [6] التي لا تفترض نموذج احتمال أساسي. وتشمل بعض من هذه الاختبارات اختبار مجموع مرتبة ويلكوكسون (Wilcoxon rank sum test) واختبار *U* لمان ويتني (Mann-Whitney *U* test) واختبار كروسكال واليس (Kruskal–Wallis test) واختبار التشغيل (the runs test) [6, 12]. توفّر الكتب الأكثر تقدماً تفاصيل عن إدارة هذه الاختبارات غير البارامترية. يجب أن لا ننسى أنه إذا تم استخدام التحليل الإحصائي الوارد في هذا الكتاب للمجتمعات الإحصائية أو العمليات التي لم يتم توزيعها بشكل طبيعي فإن النتائج قد تكون ذات قيمة قليلة وقد يفقد الباحث نتائج هامة من البيانات.

بافتراض أن بياناتنا تمثل مجتمعًا إحصائيًّا أو عملية مُوزِّعة بشكل طبيعي فإننا جاهزون عندئذ لإجراء مجتمع إحصائي متنوع من الاختبارات الإحصائية التي تسمح لنا باختبار الفروض فيما يتعلق بمساواة المتوسطات والتباينات عبر مجتمعين إحصائيين أو أكثر. لنتذكر أنه من أجل التوزيع الطبيعي هناك حاجة فقط إلى المتوسط والانحراف المعياري للتوصيف التام للطبيعة الاحتمالية للمجتمع الإحصائي أو العملية. وهكذا إذا كان علينا مقارنة مجتمعين مُوزَّعين بشكل طبيعي فإننا بحاجة فقط إلى مقارنة المتوسطات والتغيرات لهذين المجتمعين الإحصائيين. وإذا لم تكن المجتمعات الإحصائية أو العمليات مُوزَّعة بشكل طبيعي فقد تكون هناك بارامترات أخرى مثل عدم التماثل والتفرطح (kurtosis) تميز بين اثنتين أو أكثر من المجتمعات الإحصائية أو العمليات.

## (٥,١) مقارنة متوسطات المجتمعات الإحصائية
## COMPARISON OF POPULATION MEANS

إن أحد الأسئلة الأساسية التي طرحها العلماء والمهندسون الذين يقومون بعمل التجارب هو ما إذا كان مجتمعان إحصائيان أو طريقتان أو علاجان مختلفيْن حقًّا في النزعة المركزية. وبشكل أكثر تحديداً هل متوسطًا المجتمعين المنعكسين في متوسطي العينتين اللتين تم جمعهما تحت شرطين تجريبيين مختلفان اختلافاً كبيراً؟ أو هل الفرق الذي تم ملاحظته بين متوسطين هو ببساطة بسبب الصدفة وحدها؟

بعض الأمثلة أو الأسئلة التي طرحها المهندسون الطبيون الحيويون التي تتطلب المقارنة بين متوسطّي مجتمعين إحصائيين تشمل ما يلي :

١ – هل أحد أدوية العلاج الكيميائي أكثر فعالية من دواء علاج كيميائي آخر في تقليص حجم الورم السرطاني؟

٢ – هل هناك فرق في وضعية الراحة بين مجتمع الشباب ومجتمع كبار السن؟

٣- هل هناك فرق في كثافة العظام للنساء اللواتي قبل سن اليأس مقابل أولئك اللواتي بعد سن اليأس؟

٤- هل التيتانيوم مادة أقوى من الفولاذ لعملية زرع العظام؟

٥- بالنسبة للتصوير بالرنين المغناطيسي هل أداء أحد أنواع متتاليات النبضة أفضل من الآخر في الكشف عن مساحات (سُبُل) المادة البيضاء في الدماغ؟

٦- هل تمنع الدعامات داخل الأوعية الشاطفة للأدوية (drug-eluting intravascular stents) عودة التضيق على نحو أكثر فعالية من الدعامات غير المُغلَّفة (noncoated stents)؟

للإجابة على هذا النوع من الأسئلة يجمع مهندسو الطبية الحيوية في كثير من الأحيان عينات من مجتمعين مختلفتين للأشخاص أو من مجتمع واحد من الأشخاص ولكن تحت ظرفين مختلفين. وبالنسبة لقياس محدد يمثل العينات يتم عادة تقدير متوسط العينة لكل مجتمع أو تحت كل ظرف من الظرفين. ربما تكون مقارنة متوسطي مجتمعين إحصائيين منعكسين في مجموعتين من البيانات التحليل الإحصائي الأكثر ذكراً في المراجع العلمية والهندسية ويمكن إنجازه باستخدام ما يعرف بالاختبار $t$.

### (١, ١, ٥) الاختبار $t$ The t Test

عند إجراء الاختبار $t$ نطرح السؤال التالي "هل المتوسطان لمجتمعين إحصائيين مختلفان حقًّا؟" أو "هل سنرى الفروق الملحوظة فقط بسبب الصدفة العشوائية؟" إن المجتمعين منعكسان في بيانات تم جمعها تحت ظرفين مختلفين. وقد تشمل هذه الظروف علاجين أو عمليتين مختلفتين.

لمعالجة هذه المسألة نستخدم أحد الاختبارين التاليين اعتماداً على ما إذا كانت مجموعتا البيانات مستقلتين أو تعتمدان على بعضهما:

١- اختبار $t$ غير مزدوج لمجتمعين إحصائيين من البيانات المستقلة.

٢- اختبار $t$ مزدوج لمجتمعين إحصائيين من البيانات التابعة.

قبل وصف كل نوع من أنواع الاختبار $t$ نحن بحاجة إلى مناقشة فكرة اختبار الفروض.

**(١,١,١,٥) اختبار الفروض Hypothesis Testing**

كلما أجرينا التحليل الإحصائي نختبر بعض الفروض. في الواقع حتى قبل جمع البيانات نقوم بصياغة فرضية ومن ثم تصميم تجربتنا بعناية وجمع وتحليل البيانات لاختبار الفروض. إن نتيجة الاختبار الإحصائي إذا كان نموذج الاحتمال المُفترض صحيحًا تسمح لنا بقبول أو رفض الفروض وفعل ذلك مع مستوى معين من الثقة.

هناك أساساً نوعان من الفروض نختبرهما عند إجراء التحليل الإحصائي. هما فرضية العدم (الصفرية) (null hypothesis) والفروض البديلة (alternative hypothesis).

يتم التعبير عن فرضية العدم التي يُرمز لها بـ $H_0$ على النحو التالي بالنسبة للاختبار $t$ مقارناً متوسطيّ مجتمعين $\mu_1$ و $\mu_2$ :

$$H_0 : \mu_1 = \mu_2$$

يتم التعبير عن الفرضية البديلة التي يُرمز لها بـ $H_1$ على أنها أحد الاختبارات التالية بالنسبة للاختبار $t$ مقارناً متوسطيّ مجتمعين $\mu_1$ و $\mu_2$ :

$H_0 : \mu_1 \neq \mu_2$ (الاختبار $t$ ثنائي الذيل)

$H_0 : \mu_1 < \mu_2$ (الاختبار $t$ أحادي الذيل)

أو

$H_0 : \mu_1 > \mu_2$ (الاختبار $t$ أحادي الذيل)

إذا كانت $H_1$ من الشكل الأول حيث لا نعرف مسبقاً من مجتمع البيانات أي متوسط سيكون أكبر من الآخر فإننا سوف نجري الاختبار ثنائي الذيل وهو ما يعني

ببساطة أن مستوى الأهمية الذي من أجله نقبل أو نرفض فرضية العدم سيكون ضعف مستوى الحالة التي نتنبأ فيها مقدماً من التجربة أن أحد المتوسطين سيكون أقل من الثاني استناداً إلى الفسيولوجيا أو العملية الهندسية أو الظروف الأخرى التي تؤثر على النتائج التجريبية. إن الطريقة الأخرى للتعبير عن ذلك هي أنه بالاختبار أحادي الذيل سيكون لدينا ثقة في رفض أو قبول فرضية العدم أكبر منها في حالة الاختبار ثنائي الذيل.

**(٥,١,١,٢) تطبيق الاختبار $t$ Applying the $t$ Test**

الآن بعد أن حددنا فرضياتنا نحن جاهزون لإجراء الاختبار $t$. وبالأخذ في الاعتبار مجتمعين لهما $n_1$ و $n_2$ عينة نستطيع مقارنة متوسطيّ المجتمعين باستخدام الاختبار $t$. ومن المهم أن نتذكر الافتراضات الأساسية الواردة في استخدام الاختبار $t$ لمقارنة متوسطيّ المجتمعين:

١- التوزيعات الأساسية لكلا المجتمعين طبيعية.

٢- تغيرات المجتمعين متساوية تقريباً: $s_1^2 = s_2^2$.

هناك افتراضات كبيرة نقوم بها عندما تكون $n_1$ و $n_2$ صغيرة. وإذا كانت هذه الافتراضات ضعيفة بالنسبة للبيانات التي يجري تحليلها فإننا بحاجة إلى إيجاد إحصائيات مختلفة لمقارنة المجتمعين الإحصائيين.

وبالأخذ في الاعتبار مجتمعين من البيانات التي تم أخذ عيناتها $x_i$ و $y_i$ فإن المتوسطين للمجتمعين أو العمليتين المنعكستين في البيانات التي تم أخذ عيناتها يمكن مقارنتهما باستخدام الإحصاء $t$ التالي:

**(٥,١,١,٣) الاختبار $t$ غير المزدوج Unpaired $t$ Test**

يمكن تقدير الإحصاء $t$ غير المزدوج باستخدام:

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{(n_1 - 1)S_x^2 + (n_2 - 1)S_y^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

حيث $n_1$ هو عدد الملاحظات $x_i$ و $n_2$ هو عدد الملاحظات $y_i$ $S_x^2$ هو تباين العينة لـ $x_i$ $S_y^2$ هو تباين العينة لـ $y_i$ $\overline{x}$ هو متوسط العينة لـ $x_i$ و $\overline{y}$ هو متوسط العينة لـ $y_i$.

حالما يتم حساب الإحصاء $T$ يمكننا مقارنة القيمة $T$ المُقدَّرة مع القيم $t$ الواردة في جدول للتوزيع $t$. وبشكل أكثر تحديداً إذا أردنا رفض فرضية العدم بمستوى مقداره $1-\alpha$ من الثقة فإننا بحاجة عندئذ إلى تحديد ما إذا كانت القيمة $T$ المُقدَّرة أكبر من مدخل القيمة $t$ في الجدول $t$ المرتبطة بمستوى كبير لـ $\alpha$ (اختبار $t$ أحادي الجانب) أو $\alpha/2$ (اختبار $t$ ثنائي الجانب). وبعبارة أخرى نحن بحاجة إلى معرفة القيمة $t$ من التوزيع $t$ الذي تكون من أجله المساحة تحت المنحنى $t$ إلى اليمين من $t$ مساوية إلى $\alpha$. يجب أن تكون $T$ المُقدَّرة أكبر من هذه الـ $t$ لرفض فرضية العدم بمستوى $1-\alpha$ من الثقة.

وهكذا نحن نقارن القيمة $T$ مع مدخل جدول التوزيع $t$ من أجل:

$t(\alpha, n_1+n_2-2)$ (أحادي الجانب)

أو

$t(\alpha/2, n_1+n_2-2)$ (ثنائي الجانب)

حيث $\alpha$ هو مستوى الأهمية (يساوي 1− مستوى الثقة) و $n_1$ و $n_2$ هما عدد العينتين من المجتمعين اللذين تُجري مقارنتهما.

لاحظ أن $\alpha$ هو مستوى الأهمية الذي نريد عنده قبول أو رفض فرضيتنا. نرفض فرضية العدم بالنسبة لمعظم البحوث عندما $\alpha \le 0.05$. وهذا يتوافق مع مستوى ثقة مقداره 95%.

على سبيل المثال إذا أردنا رفض فرضية العدم بمستوى ثقة مقدارها 95% فإنه عندئذ بالنسبة لاختبار $t$ أحادي الجانب يجب أن تكون $T$ المُقدَّرة أكبر من $t(0.05, n_1+n_2-2)$ لرفض $H_0$ وقبول $H_1$ بثقة مقدارها 95% أو مستوى أهمية تساوي 0.05 أو أقل. علينا أن نتذكر أن ثقتنا = $(1-0.05)\times 100\,\%$. وكان هذا الاختبار من أجل $H_1: \mu_1 < \mu_2$ أو $H_1: \mu_1 > \mu_2$

(أحادي الجانب). إذا كانت $H_1$ هي $H_1 : \mu_1 \neq \mu_2$ فإن $T$ المُقاسة يجب أن تكون عندئذ أكبر من $t(0.05/2, n_1 + n_2 - 2)$ لرفض فرضية العدم بنفس الثقة 95%.

وهكذا لاختبار الفرضية البديلة التالية :

١- بالنسبة لـ $H_1 : \mu_1 \neq \mu_2$ : استخدم اختبار $t$ ثنائي الذيل. بالنسبة لـ $\alpha < 0.05$ فإن $T > t(0.025, n_1 + n_2 - 2)$ المُقدرة لرفض $H_0$ بثقة مقدارها 95%.

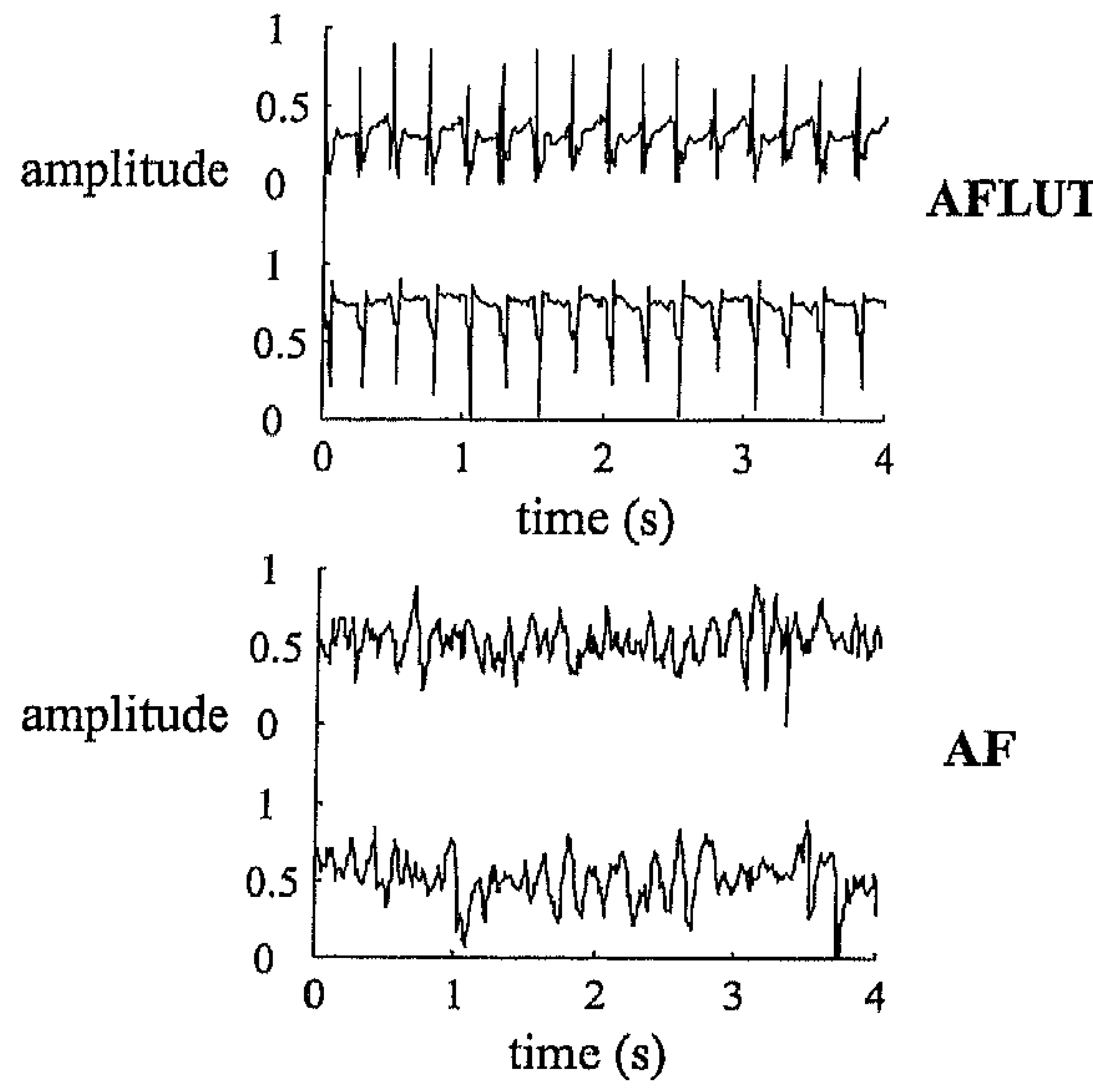
٢- بالنسبة لـ $H_1 : \mu_1 < \mu_2$ أو $H_1 : \mu_1 > \mu_2$ : استخدم اختبار $t$ أحادي الذيل. بالنسبة لـ $\alpha < 0.05$ فإن $T > t(0.05, n_1 + n_2 - 2)$ المُقدرة لرفض $H_0$ بثقة مقدارها 95%.

يُشار أيضاً إلى $\alpha$ بأنها نوع الخطأ الأول (I). بالنسبة للاختبار $t$ فإن $\alpha$ هي احتمال ملاحظة قيمة $t$ مُقاسة أكبر من إدخال الجدول $t(\alpha; df)$ إذا كان المتوسطان الحقيقيان لمجتمعين $x$ و $y$ متساويين فعلاً. وبعبارة أخرى لا يوجد فرق كبير $(\alpha < 0.05)$ في متوسطيّ المجتمعين ولكن تحليل البيانات التي تم أخذ عيناتها أدت بنا إلى الاستنتاج بأن هناك فرقاً وبالتالي رفض فرضية العدم. يُشار إلى هذا بأنه نوع الخطأ الأول.

**مثال (٥,١)**

إن أحد الأمثلة على تحدٍّ في الهندسة الطبية الحيوية حيث يمكننا استخدام الاختبار $t$ لتحسين عمل الأجهزة الطبية هو تطوير أجهزة إزالة الرجفان القابلة للزرع لاكتشاف وعلاج نَظم القلب غير الطبيعي [13]. تعد أجهزة إزالة الرجفان القابلة للزرع أجهزة إلكترونية صغيرة يتم وضعها في الصدر ولها أسلاك رفيعة يتم وضعها في حجرات القلب. تحتوي هذه الأسلاك على إلكترودات تكشف التيار الكهربائي الصغير العابر لعضلة القلب. تتبع هذه التيارات في ظل الظروف العادية نمطاً منظماً جداً للتوصيل ضمن عضلة القلب. يبين الشكل (٥,١) مثالاً على مخطط كهربائي أو تسجيل لنشاط كهربائي يتم تحسسه بواسطة أحد الإلكترودات في ظل ظروف طبيعية. عندما لا يعمل القلب بشكل طبيعي ويدخل في حالة رجفان (الشكل ٥,١) حيث لا يتقلص

القلب بشكل طبيعي أو يضخ الدم إلى بقية أجزاء الجسم ينبغي للجهاز صدم القلب بتيار كهربائي كبير من الجهاز في محاولة لإعادة نَظم القلب إلى طبيعته. ومن المهم لعدد من الأسباب أن يكشف الجهاز بدقة بداية اضطرابات النّظم المهددة للحياة مثل الرجفان وإعطاء الصدمة المناسبة. لإعطاء الصدمات يجب على الجهاز استخدام بعض أنواع خوارزميات معالجة الإشارات لكي يحدد بشكل آلي أن المخطط الكهربائي غير طبيعي وأنه صفة للرجفان.

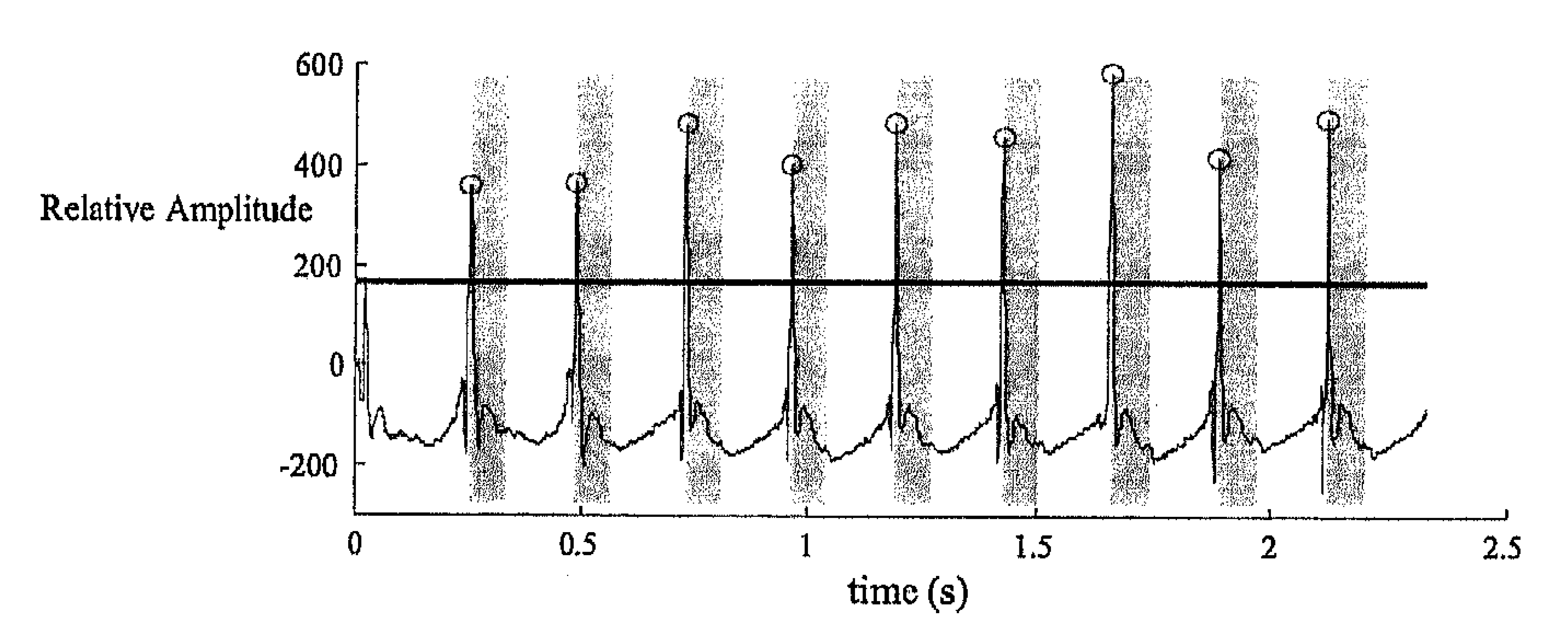



**الشكل (١ ,٥). تسجيلات مخطط كهربائي مُقاسة بواسطة إلكترودات موضوعة داخل الأذين الأيسر للقلب. وبالنسبة لكل عملية نَظم هناك تسجيلان للمخطط الكهربائي مأخوذان من موقعين مختلفين في الأذين: الرفرفة الأذينية (AFLUT) والرجفان الأذيني (AF).**

إن إحدى الخوارزميات المُستخدمة في معظم الأجهزة لتمييز نَظم القلب الطبيعي عن الرجفان هي خوارزمية المعدل. هذه الخوارزمية هي في الأساس خوارزمية عبور عتبة مطال يحدد الجهاز بواسطتها عدد المرات التي يتجاوز فيها المخطط الكهربائي عتبة

المطال في فترة محددة من الزمن ومن ثم يقدر المعدل من عدد المرات المُكتشفة لعبور العتبة. ويوضح الشكل (٥,٢) كيف تعمل هذه الخوارزمية.

قبل وضع مثل هذه الخوارزمية في أجهزة إزالة الرجفان القابلة للزرع وجب على الباحثين والمطورّين إثبات ما إذا كان المعدل المُقدَّر من قبل مثل هذا الجهاز مختلفاً فعلاً بين عمليات النَظم الطبيعية والرجفانية (fibrillatory). وكان من المهم أن يكون لديها تداخل بسيط بين عمليات النَظم الطبيعية والرجفانية بحيث يمكن تحديد عتبة المعدل حيث إن المعدلات التي تتجاوز العتبة قد تؤدي بالجهاز إلى إعطاء صدمة. ولكي تعمل هذه الخوارزمية يجب أن يكون هناك اختلاف كبير في المعدلات بين عمليات النَظم الطبيعية والرجفانية وتداخل بسيط من شأنه أن يؤدي إلى صدمات يتم إعطاؤها بشكل غير صحيح والتسبب في آلام شديدة أو خطر تحريض الرجفان. يمكن للتداخل في المعدلات بين عمليات النظم الطبيعية والرجفانية أن ينتج أيضاً عن الجهاز الذي لا يكشف الرجفان بسبب المعدلات المنخفضة.



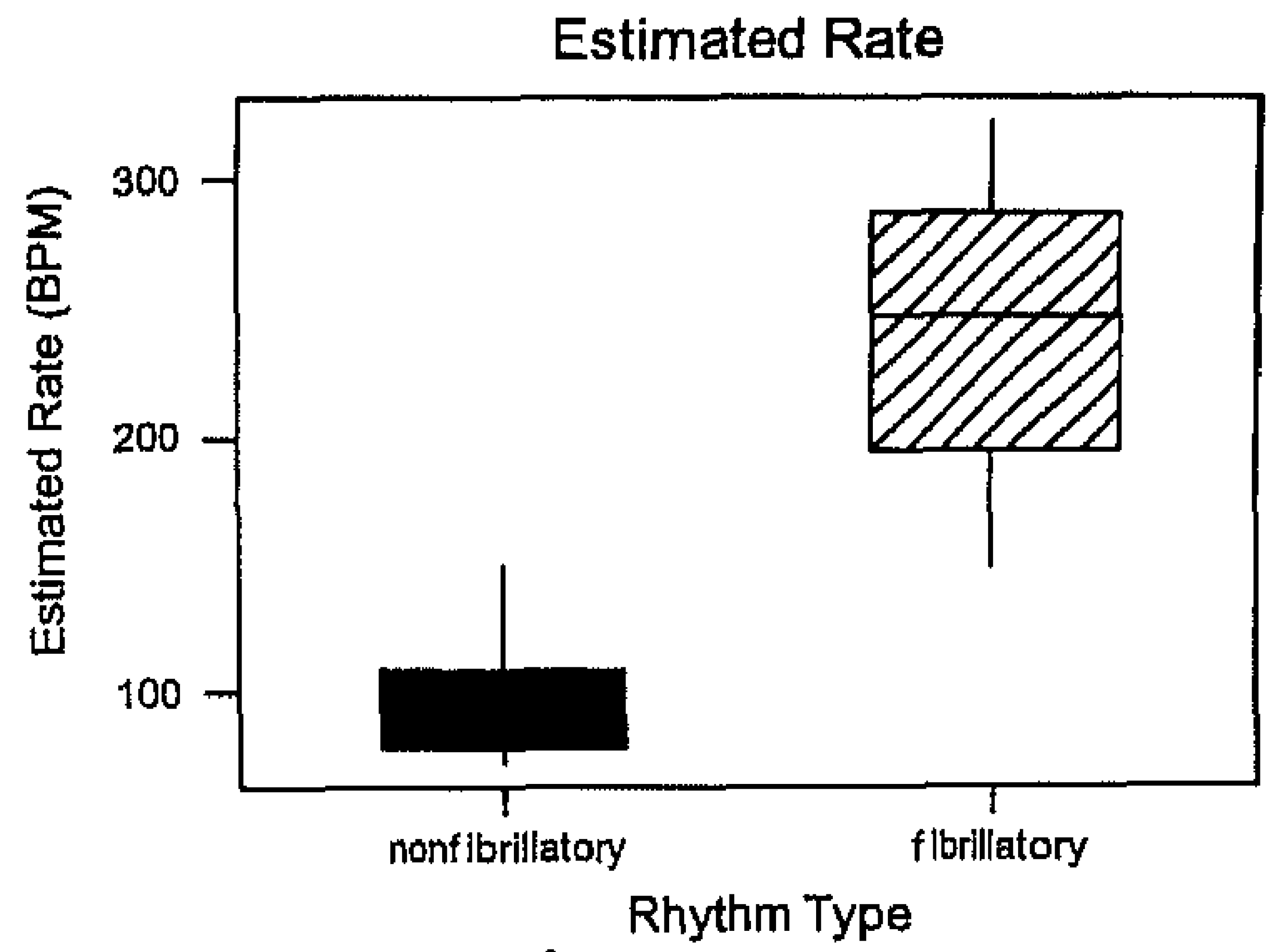

الشكل (٢ , ٥). معدَّل ضربات قلب مُقدَّر من مخطط كهربائي باستخدام خوارزمية عتبة المطال. كلما تجاوز مطال إشارة المخطط الكهربائي (شكل الموجة المستمر) عتبة المطال (الخط الأفقي الرمادي المتصل) خلال فترة زمنية محددة (الأشرطة العمودية المُظللة) يتم الكشف عن "حدث". يتم حساب المعدل عن طريق حساب عدد الأحداث في فترة زمنية معينة.

لتحديد ما إذا كان المعدل خوارزمية جيدة للكشف عن النَظم الرجفاني قد يسجل الباحثون في الواقع إشارات المخطط الكهربائي من مرضى حقيقيين أظهروا نَظمًا طبيعيًّا أو رجفانيًّا وقد يستخدمون خوارزمية المعدل لتقدير المعدل ومن ثم يقارنون معدلات المتوسط للنَظم الطبيعي مقابل معدلات المتوسط للنَظم الرجفاني. على سبيل المثال لنفترض أن الباحثين جمعوا تسجيلات مخطط كهربائي مدتها 15 ثانية للحصول على أمثلة للرجفان $(n_1 = 10)$ والنَظم عديم الرجفان $(n_2 = 11)$ في 21 مريضًا. نلاحظ أنه تم جمع بيانات الرجفان من أشخاص مختلفين عن البيانات العادية. وهكذا لم يكن لدينا حجب في هذا التصميم التجريبي.

تم تقدير المعدل وذلك باستخدام خوارزمية الجهاز لكل تسجيل مدته 15 ثانية للفرد. يبين الشكل (٥,٣) مخطط الصندوق والمؤشر لكل من مجموعتي البيانات.

ويُعطى الإحصاء الوصفي لمجموعتي البيانات حسب الجدول التالي:

| | المخططات الكهربائية عديمة الرجفان $(N_1=10)$ | المخططات الكهربائية الرجفانية $(N_1=11)$ |
|---|---|---|
| المتوسط | 96.82 | 239.0 |
| الانحراف المعياري | 22.25 | 55.3 |




الشكل (٣, ٥). مخطط الصندوق والمؤشر لبيانات المعدل المُقدَّرة من عينات نَظم قلب عديم الرجفان ونَظم قلب رجفاني.

هناك عدة أشياء يجب ملاحظتها من المخططات والإحصاء الوصفي. أولاً إن حجم العينة صغير وبالتالي فمن المرجح أن التغيُّر لتسجيلات المخطط الكهربائي من النَظم الطبيعي والرجفاني لا يتم تمثيله بشكل مناسب. وعلاوة على ذلك فإن البيانات غير متماثلة ولا تظهر بأنها مُنمذجة بشكل جيد بواسطة التوزيع الطبيعي. وأخيراً يبدو أن التغير في المعدلات للنَظم الرجفاني أكبر بكثير منه للنَظم الطبيعي. وهكذا فإن بعض الافتراضات التي نقوم بها بخصوص تسوية المجتمعات الإحصائية والمساواة بين التباينات ربما تكون ناقصة في تطبيق الاختبار $t$. ولكن لغرض التوضيح فإننا سوف نجري الاختبار $t$ باستخدام الإحصاء $t$ غير المزدوج.

ولإيجاد الإحصاء $T$ المُقدَّر من بيانات العينة فإننا نستبدل المتوسطات والتباينات وحجم العينة في معادلة الإحصاء $T$ غير المزدوج. وبالنسبة لهذا المثال نجد أن قيمة $t$ المُقدَّرة هي 7.59. وإذا نظرنا إلى جداول $t$ من أجل [11+10−2] درجات حرية نستطيع أن نرفض $H_0$ عندما $\alpha<0.005$ لأن قيمة $T$ المُقدَّرة تتجاوز إدخالات جدول $t$ عندما $\alpha=0.05\,(t=1.729)$ و $\alpha=0.01\,(t=2.539)$ و $\alpha=0.005\,(t=2.861)$.

وهكذا فإننا نرفض فرضية العدم بثقة مقدارها $(1-0.005)\times 100\%$ وندرك أن متوسط معدل ضربات القلب للنَظم الرجفاني أكبر من متوسط معدل ضربات القلب للنَظم الطبيعي. وهكذا ينبغي لخوارزمية المعدل العمل جيداً في التمييز بين النُظم الطبيعي والنَظم الرجفاني. ومع ذلك فقد اختبرنا متوسطات المجتمعات الإحصائية فقط. كما يلاحظ المرء في الشكل (٥،٣) هناك بعض التداخل في العينات الفردية بين النَظم الطبيعي والرجفاني. وهكذا يمكننا أن نتوقع بأن الجهاز سيقوم بأخطاء في إعطاء صدمة بشكل غير صحيح عندما يكون القلب في حالة نَظم طبيعي ولكن متسارع (كما في التمارين) أو قد يفشل الجهاز في إعطاء صدمة عندما يكون القلب مرتجفاً ولكن بمعدل بطيء أو بمطال منخفض.

عند تطبيق الاختبار $t$ فمن المهم أن نلاحظ أنه لا يمكن أبداً إثبات أن متوسطَين متساويان. تستطيع الإحصائيات فقط إظهار أن اختباراً محدداً لا يمكنه إيجاد فرق في متوسطات المجتمعات وعدم إيجاد فرق لا يعني إثبات المساواة. إن فرضية العدم أو عدم الوجود هي أنه لا يوجد فرق في المتوسطات بغض النظر عن ما هو الفرق الحقيقي بين المتوسطَين. إن عدم إيجاد فرق بالبيانات التي تم جمعها والاختبار الإحصائي المناسب لا يعني إثبات أن المتوسطات متساوية. وهكذا فإننا ببساطة نقبل فرضية العدم ولا نقبل بمستوى الثقة أو الأهمية. إننا نحدد مستوى الثقة فقط عندما نرفض فرضية العدم.

**(٤, ١, ١, ٥) الاختبار $t$ المزدوج Paired $t$ Test**

استخدمنا في المثال السابق الاختبار $t$ غير المزدوج لأن مجموعتي البيانات جاءتا من مجتمعين مختلفين غير مترابطين للمرضى. تكمن المشكلة في مثل هذا التصميم التجريبي في أن الاختلافات في مجتمعي المرضى قد تؤدي إلى اختلافات في متوسط معدل ضربات القلب الذي لا علاقة له بنَظم القلب الفعلي وإنما بالاختلافات في حجم أو عمر القلوب بين مجموعتي المرضى أو بعض الاختلافات الأخرى بين مجموعات المرضى. هناك طريقة أفضل لإجراء الدراسة السابقة وهي جمع بيانات القلب الطبيعية والرجفانية من مجتمع واحد من المرضى. وسوف يحتاج فقط أحد أجهزة الرجفان إلى التمييز بين النَظم الطبيعي والنَظم الرجفاني ضمن مريض واحد. يمكننا من خلال حجب أشخاص التخلص من التغير بين الأشخاص في خصائص المخطط الكهربائي التي يمكن أن تعيق خوارزمية المعدل من المجتمعات الإحصائية المنفصلة. قد يكون من المعقول افتراض أن معدلات النَظم الطبيعي والنَظم الرجفاني تختلف داخل المريض أكثر منها عبر المرضى. وبعبارة أخرى نجمع بالنسبة لكل مريض مخططاً كهربائياً خلال عمل القلب الطبيعي وخلال الرجفان. في مثل هذا التصميم التجريبي قد نقارن المتوسطيْن لمجموعتي البيانات باستخدام اختبار $t$ المزدوج.

نستخدم الاختبار $t$ المزدوج عندما يكون هناك ازدواج طبيعي بين كل ملاحظة في مجتمع البيانات $X$ مع كل ملاحظة في مجتمع البيانات $Y$. على سبيل المثال قد يكون لدينا الفروض التالية التي تستدعي الاختبار $t$ المزدوج :

١- جمع قياسات ضغط الدم من 8 مرضى قبل وبعد ممارسة التمرين.

٢- جمع كل من بيانات الـ MR (الرنين المغناطيسي) والـ CT (التصوير المقطعي المحوسب) من 20 مريض لمقارنة جودة صور الأوعية الدموية.

٣- جمع زمن الحوسبة لخوارزمية معالجة صورة قبل وبعد حدوث تغيير في البرنامج.

في مثل هذه الحالات لم تعد مجموعات البيانات $X$ و $Y$ مستقلة. في المثال الأول أعلاه إن الفسيولوجيا والبيئة البيولوجية التي تؤثر على ضغط الدم قبل التمرين في كل مريض تؤثر أيضاً على ضغط الدم بعد التمرين وذلك لأننا نجمع البيانات قبل وبعد التمرين من نفس الوحدات التجريبية. وهكذا فإن هناك عددًا من المتغيرات التي تؤثر على ضغط الدم لا نستطيع التحكم بها مباشرة ولكن يمكن التحكم بآثارها على ضغط الدم (إلى جانب تأثير التمرين) باستخدام نفس الوحدات التجريبية (الكائنات البشرية في هذه الحالة) لكل مجتمع بيانات. في مثل هذه الحالات تكون الوحدات التجريبية بمثابة تحكم ذاتي بنفسها. وهذا هو عادة التصميم التجريبي المُفضل لمقارنة المتوسطات.

وبالنسبة للاختبار $t$ المزدوج لدينا مرة أخرى فرضية العدم والفروض البديلة كما ذُكر أعلاه للاختبار $t$ غير المزدوج. ومع ذلك فإننا نستخدم في الاختبار $t$ المزدوج الاختبار $t$ على الفرق $W_i = X_i - Y_i$ بين نقاط البيانات المزدوجة من كل مجتمع من المجتمعين الإحصائيين.

نحسب الآن الإحصائية $T$ المزدوجة : ($n$ = عدد الأزواج)

$$T = \frac{\overline{W}}{S_w / \sqrt{n}}$$

حيث $\overline{W}$ هو متوسط الفرق للاختلافات $W_i$ و $S_w$ هو الانحراف المعياري للاختلافات $W_i$.

وكما هو بالنسبة للاختبار $t$ غير المزدوج لدينا الآن قيمة $T$ المُقدرة التي نستطيع مقارنتها مع قيم $t$ في جدول توزيع $t$ لتحديد ما إذا كانت القيمة $T$ المُقدَّرة تقع ضمن القيم القصوى (أكبر من 95%) للتوزيع $t$.

ولرفض فرضية العدم عند مستوى أهمية مقدارها $\alpha$ (مستوى ثقة مقدارها $1-\alpha$) يجب أن تكون قيمة $T$ المُقدَّرة أكبر من $t(\alpha,n-1)$ حيث $n$ هو عدد أزواج البيانات.

إذا كانت قيمة $T$ المُقدَّرة تتجاوز $t(\alpha,n-1)$ فإننا نرفض $H_0$ ونقبل $H_1$ عند مستوى أهمية يساوي $\alpha$ أو مستوى ثقة مقدارها $1-\alpha$.

مرة أخرى إذا كانت $H_1 : \mu_1 < \mu_2$ أو $H_1 : \mu_1 > \mu_2$ فإننا نجري اختباراً أحادي الجانب حيث يجب أن تكون الإحصائية $T$ أكبر من $t(\alpha,n-1)$ لرفض فرضية العدم عند مستوى مقداره $\alpha$. إذا كانت فرضية العدم التي نبدأ بها قبل التجربة $H_i : (\mu_1 \neq \mu_2)$ فإننا نجري عندئذ اختباراً ثنائي الجانب ويجب أن تكون الإحصائية $T$ أكبر من $t(\alpha/2\,,n-1)$ لرفض فرضية العدم عند مستوى أهمية يساوي $\alpha$.

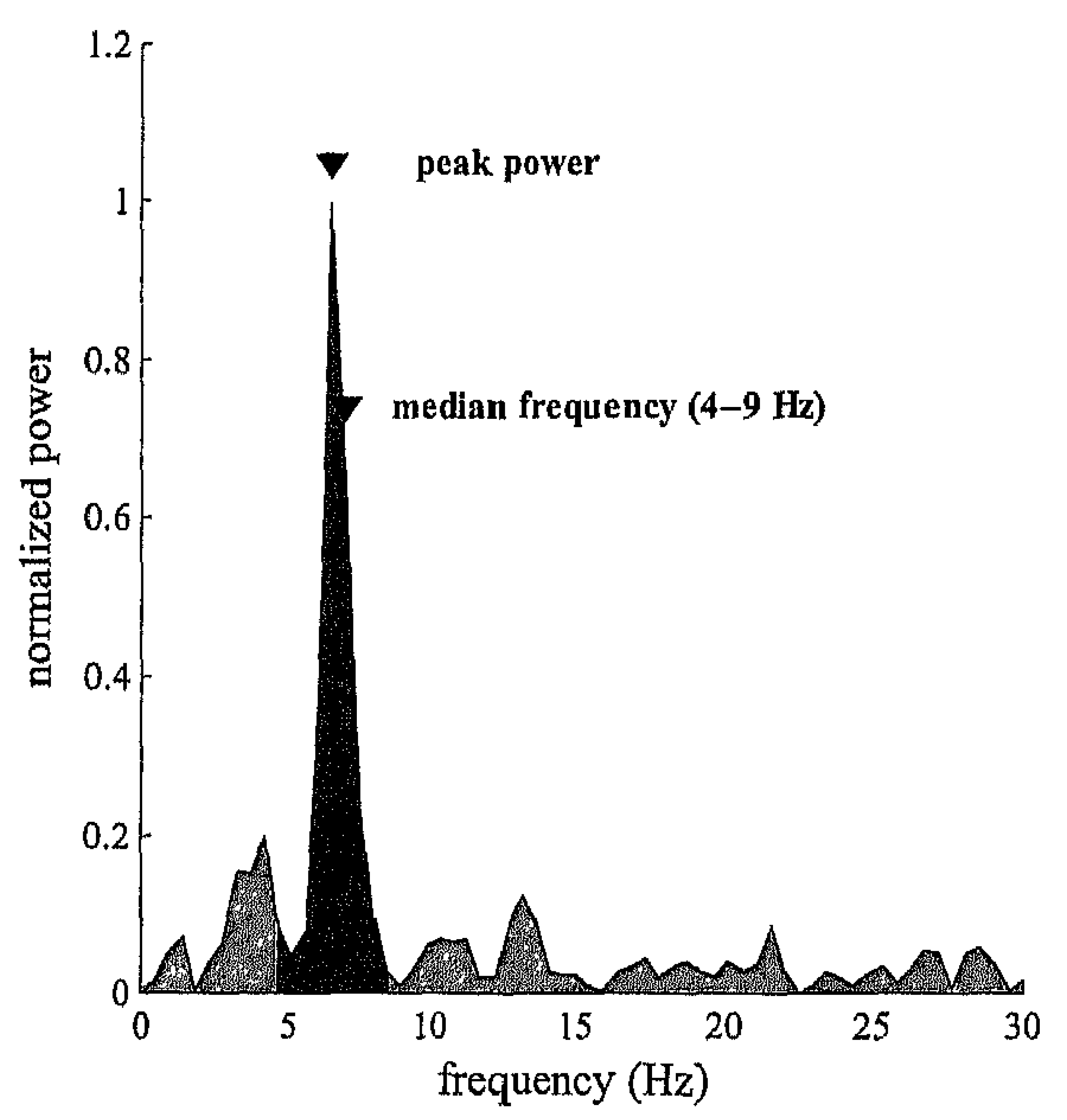
**(٥,١,١,٥) مثال لتحدٍّ في الهندسة الطبية الحيوية**

**Example of a Biomedical Engineering Challenge**

فيما يتعلق بالنَظم غير الطبيعي للقلب الذي تم مناقشته سابقاً يمكن استخدام أدوية مضادة لاضطراب النَّظْم لإبطاء النَّظم غير الطبيعي أو إنهائه مثل الرجفان. على سبيل المثال قد يتم استخدام دواء مثل بروكايناميد (procainamide) لإنهاء الرجفان الأذيني. إن الآلية الدقيقة التي يؤدي بواسطتها الدواء إلى إنهاء النظم غير معروفة بالضبط ولكن يُعتقد أن الدواء يقوم بتغيير فترة الاستعصاء وسرعة التوصيل لخلايا القلب [8]. غالباً ما يستخدم المهندسون الطبيون الحيويون معالجة الإشارة على الإشارات الكهربائية المتولدة بواسطة القلب كوسائل غير مباشرة لدراسة الفسيولوجيا (علم وظائف الأعضاء) الأساسية. وبشكل أكثر تحديداً فقد يستخدم المهندسون التحليل الطيفي أو تحليل فورييه لدراسة التغيرات في المحتوى الطيفي للإشارة الكهربائية

مع مرور الزمن مثل المخطط الكهربائي. قد تخبرنا مثل هذه التغيرات شيئاً عن الفسيولوجيا الكهربائية الأساسية.

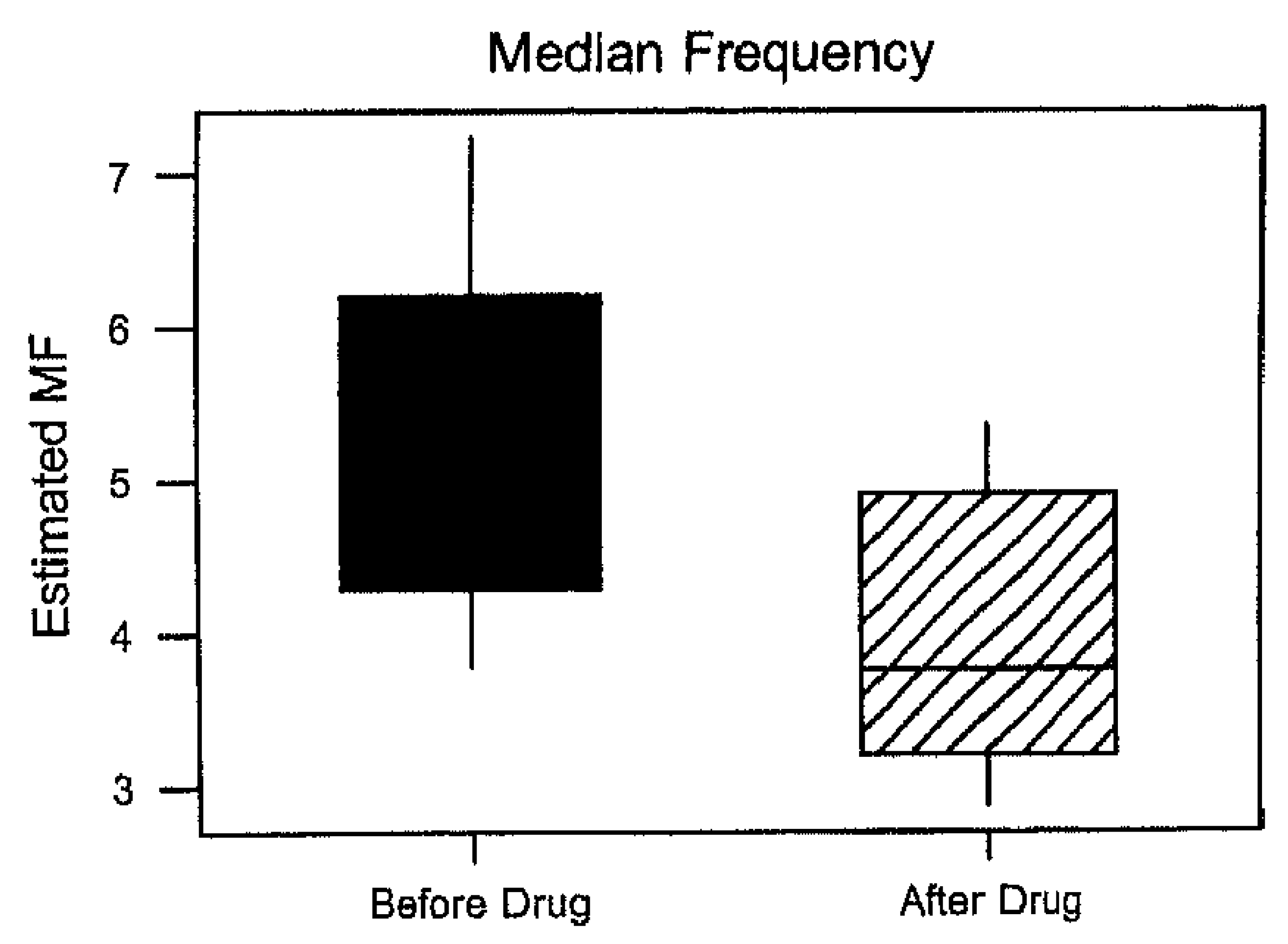
على سبيل المثال فقد تم استخدام التحليل الطيفي لدراسة التغيرات في الطيف الترددي للرجفان الأذيني مع إعطاء الدواء. في إحدى الدراسات [8] كان المهندسون الطبيون الحيويون مهتمين بدراسة التغيرات في التردد الوسط للمخططات الكهربائية الأذينية بعد إعطاء الدواء. يبين الشكل (٥,٤) مثالاً للطيف الترددي للرجفان الأذيني وموقع التردد الوسط الذي هو التردد الذي يقسم قدرة الطيف (المنطقة تحت المنحنى الطيفي بين 4 و 9 هرتز) في النصف. أحد الأسئلة الذي طرحه الباحثون هو ما إذا كان التردد الوسط ينخفض بعد إعطاء دواء مثل البروكايناميد الذي يُعتقد أنه يقوم بإبطاء النشاط الكهربائي لخلايا القلب.




**الشكل (٤ , ٥). الطيف الترددي لمثال على الرجفان الأذيني. يتم تعريف التردد الوسط بأنه ذلك التردد الذي يقسم المنطقة تحت منحنى القدرة في الحزمة من 4 - 9 هرتز في النصف.**

وبالتالي تم إجراء تجربة لتحديد ما إذا كان هناك اختلاف كبير في متوسط التردد الوسط بين الرجفان قبل إعطاء الدواء والرجفان بعد إعطاء الدواء. تم جمع مخططات كهربائية في الأذين الأيمن في 11 مريضًا قبل وبعد إعطاء الدواء. وتم تقييم تسجيلات مدتها خمس عشرة ثانية لطيف التردد وتم تقدير التردد الوسط في حزمة التردد من 4 – 9 هرتز قبل وبعد إعطاء الدواء.

يوضح الشكل (٥,٥) إحصائيات موجزة عن التردد الوسيط قبل وبعد إعطاء الدواء. والسؤال هو ما إذا كان هناك انخفاض ملحوظ في التردد الوسيط بعد إعطاء البروكايناميد. في هذه الحالة فإن فرضية العدم هي أنه لا يوجد أي تغيير في متوسط التردد الوسط بعد إعطاء الدواء. الفروض البديلة هي أن متوسط التردد الوسط ينخفض بعد إعطاء الدواء. وهكذا فإننا نقارن متوسطين وقد تم جمع مجموعتي البيانات من إحدى مجموعات المرضى. وسوف نحتاج إلى الاختبار $t$ المزدوج لرفض أو قبول فرضية العدم.




**الشكل (٥ , ٥). مخطط الصندوق والمؤشر لتردد وسط تم تقديره من عينات للرجفان الأذيني تم تسجيلها قبل وبعد إعطاء الدواء.**

لإجراء الاختبار $t$ المزدوج ننشئ عمودًا آخر $W_i$ كما ورد في الجدول التالي. إننا نجـد ما يلي لـ $W_i$ : $\overline{W}=1.262$ و $S_w=0.402$. وإذا استخـدمنا هـذه التقديرات لـ $\overline{W}$ و $S_w$ في معادلة إحصائية $T$ المزدوجة فإننا نجد أن $T=10.42$ عندما $n=11$ زوج من البيانات. وإذا قارنا قيمة $T$ المُقدَّرة مع التوزيع $t$ نجد أن قيمة $T$ أكبر من إدخال الجدول عندما $t(0.005,11-1)$ ؛ وبالتالي فقد نرفض $H_0$ عند أهمية أقل من 0.005. وبعبارة أخرى فإننا نرفض فرضية العدم عند مستوى الثقة $[1-0.005]\times100\%$.

| قبل الدواء | بعد الدواء | WI |
|---|---|---|
| 4.3 | 2.90 | 1.4 |
| 4.15 | 2.97 | 1.18 |
| 3.8 | 3.20 | 0.60 |
| 5.10 | 3.30 | 1.80 |
| 4.30 | 3.75 | 0.55 |
| 7.20 | 5.35 | 1.85 |
| 6.40 | 5.10 | 1.30 |
| 6.20 | 4.90 | 1.30 |
| 6.10 | 4.80 | 1.30 |
| 5.00 | 3.70 | 1.30 |
| 5.80 | 4.50 | 1.30 |

$W_I$ هو الانخفاض في تردد الوسيط بعد إعطاء البروكايناميد

## أخطاء في استخلاص الاستنتاجات من الاختبارات الإحصائية

عندما نجري تحليلاً إحصائياً مثل الاختبار $t$ فقد نكون مخطئين في رفض أو قبول فرضية العدم. وعندما نستخلص استنتاجات من التحليل الإحصائي فإننا نحدد عادة مستوى ثقة لاستنتاجاتنا. وهذا يعني أن هناك دائماً احتمالا إلى حد ما بأن استنتاجاتنا غير صحيحة.

هناك نوعان من الأخطاء التي قد تحدث عند استخلاص استنتاجات من التحليل الإحصائي. ويُشار إلى هذه الأخطاء بالأنواع I وII.

أخطاء النوع I:

١– يُشار إليها أيضاً بأنها الخطأ الموجب الخاطئ.

٢– تحدث عندما نقبل $H_1$ عندما تكون $H_0$ صحيحة.

٣– قد تؤدي إلى تشخيص خاطئ للمرض.

أخطاء النوع II:

١– يُشار إليها أيضاً بأنها الخطأ السالب الخاطئ.

٢– تحدث عندما نقبل $H_0$ عندما تكون $H_1$ صحيحة.

٣– قد تؤدي إلى عدم التشخيص (غالباً ما تكون أكثر خطورة من خطأ النوع I).

إذا فكرنا في البيئة الطبية فقد يحدث خطأ النوع I عندما يُعطى شخص اختبارًا تشخيصيًّا للكشف عن البكتيريا العقدية والاختبار يشير إلى أن الشخص لديه البكتيريا العقدية عندما لا يكون لدى الشخص هذه البكتيريا في الواقع. إن نتيجة مثل هذا الخطأ تعني أن الشخص ينفق المال على المضادات الحيوية التي لا تخدم أي غرض.

قد يحدث خطأ النوع II عندما يكون بالفعل لدى الشخص نفسه بكتيريا عقدية ولكن الاختبار التشخيصي يعطي نتيجة سلبية ونخلص إلى أن الشخص ليس لديه بكتيريا عقدية. في هذا المثال يكون هذا النوع من الخطأ أكثر خطورة من خطأ النوع I لأن البكتيريا العقدية التي تم تركها دون علاج يمكن أن تؤدي إلى عدد من المضاعفات الخطيرة للجسم.

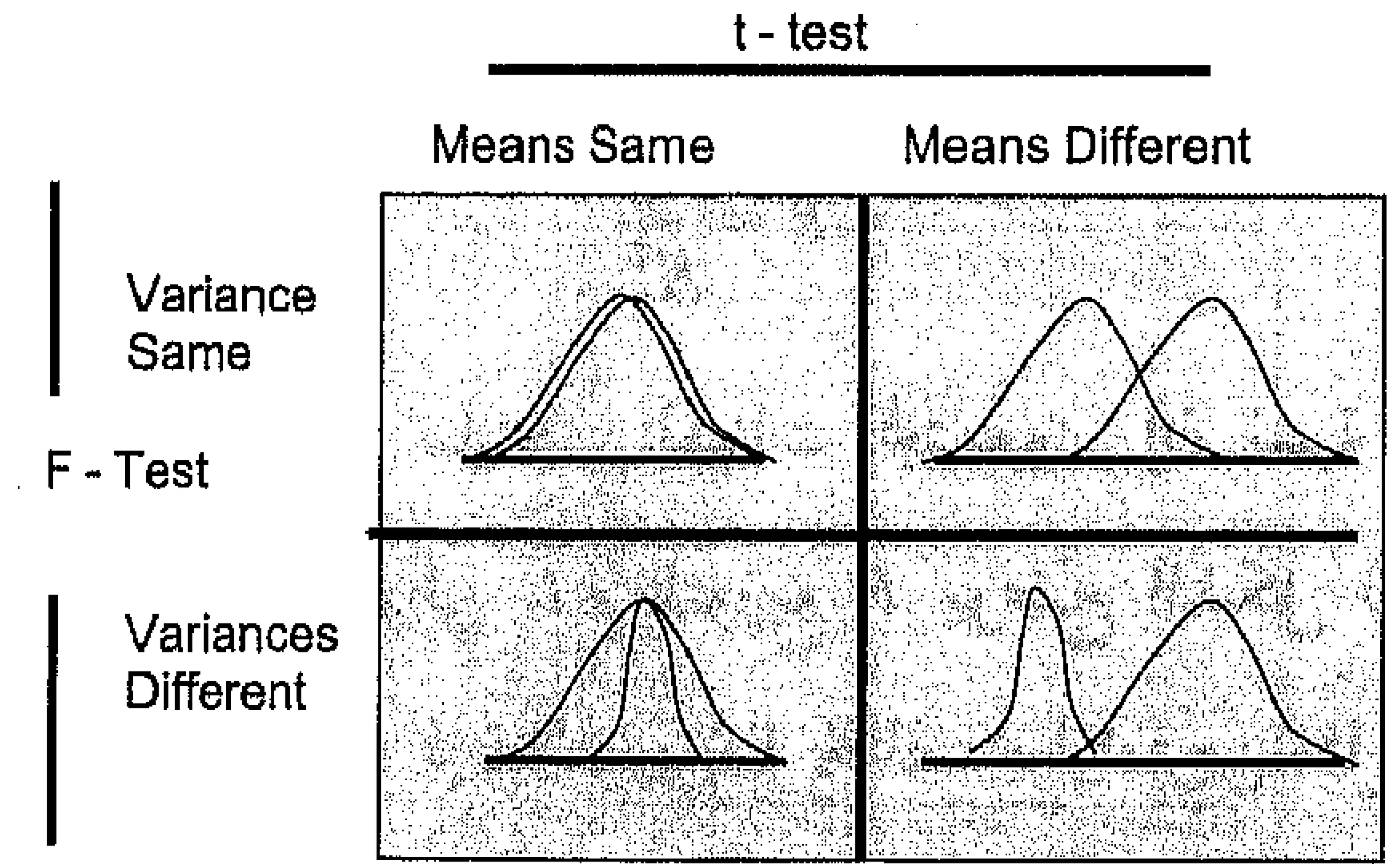
إن قيمة $\alpha$ التي نشير إليها بأنها مستوى الأهمية هي أيضاً الاحتمال لخطأ النوع I. وهكذا كلما كان مستوى الأهمية الذي قد نرفض عنده فرضية العدم أصغر كان خطأ النوع I أصغر وكان احتمال القيام بخطأ النوع I أقل.

نستخدم عادة $\beta$ للدلالة على خطأ النوع II. وسوف نناقش هذا الخطأ مرة أخرى في نهاية الفصل السابع عندما نناقش اختبارات القدرة.

## (٥,٢) مقارنة تباينين
## COMPARISON OF TWO VARIANCE

استخدمنا الاختبار $t$ لمقارنة المتوسطين لمجتمعين أو عمليتين. ويمكن أيضاً مقارنة مجتمعين فيما يتعلق بالاختلافات في التباين. وكما تمت المناقشة سابقاً يتم تماماً تمييز المجتمعات المُوزعة بشكل طبيعي بواسطة متوسطاتها وتبايناتها. وهكذا إذا أردنا الاختبار فيما يتعلق بالاختلافات بين مجتمعين طبيعيين فإننا بحاجة فقط إلى مقارنة المتوسطين والتباينين لهذين المجتمعين.

يوضح الشكل (٥,٦) دوال كثافة الاحتمال لمجتمعين طبيعيين (الرسومات السوداء والحمراء). توضح المخططات الأربعة كيف يمكن مقارنة مجتمعين مختلفين موزعين بشكل طبيعي مع بعضهما البعض. تختلف لوحتا اليمين عن لوحتي اليسار في متوسطات المجتمعين الإحصائيين. وتختلف اللوحتان العلويتان عن اللوحتين السفليتين في التباين للمجتمعات الإحصائية.




**الشكل (٥,٦). قد يختلف مجتمعان إحصائيان طبيعيان في متوسطاتها (الصف العلوي) أو تبايناتها (النصف الأيسر) أو كليهما (الزاوية اليمنى السفلى). يمكن استخدام الاختبارين t وF لاختبار الاختلافات الكبيرة في متوسطات المجتمعات الإحصائية وتباينات المجتمعات على التوالي.**

وكما هو مبين في الجزء العلوي من الرسومات يتم استخدام الاختبار $t$ لاختبار الاختلافات في المتوسط بين المجتمعين. وكما هو مبين على طول الاتجاه العمودي يتم استخدام الاختبار $F$ لاختبار الاختلافات الكبيرة في تباينات المجتمعات الإحصائية. لاحظ أن المجتمعين الطبيعيين قد يختلفان اختلافاً كبيراً في كل من المتوسط والتباين.

لمقارنة التباينات لمجتمعين نستخدم ما يُشار إليه بأنه الاختبار $F$. وكما هو بالنسبة للاختبار $t$ يفترض الاختبار $F$ أن البيانات تتألف من عينات عشوائية مستقلة من كل مجتمع من المجتمعين الطبيعيين. إذا كان المجتمعان غير موزعين بشكل طبيعي فقد تكون نتائج الاختبار $F$ لا معنى لها.

كما هو الحال مع الاختبار $t$ يتم استخدام الاختبار $F$ لاختبار الفروض التالية :

١- فرضية العدم : $H_0 : \sigma_1^2 = \sigma_2^2$ و

٢- الفروض البديلة : $H_1 : \sigma_1^2 \geq \sigma_2^2$

حيث $\sigma_1^2$ و $\sigma_2^2$ هي التباينات للمجتمعين الإحصائيين.

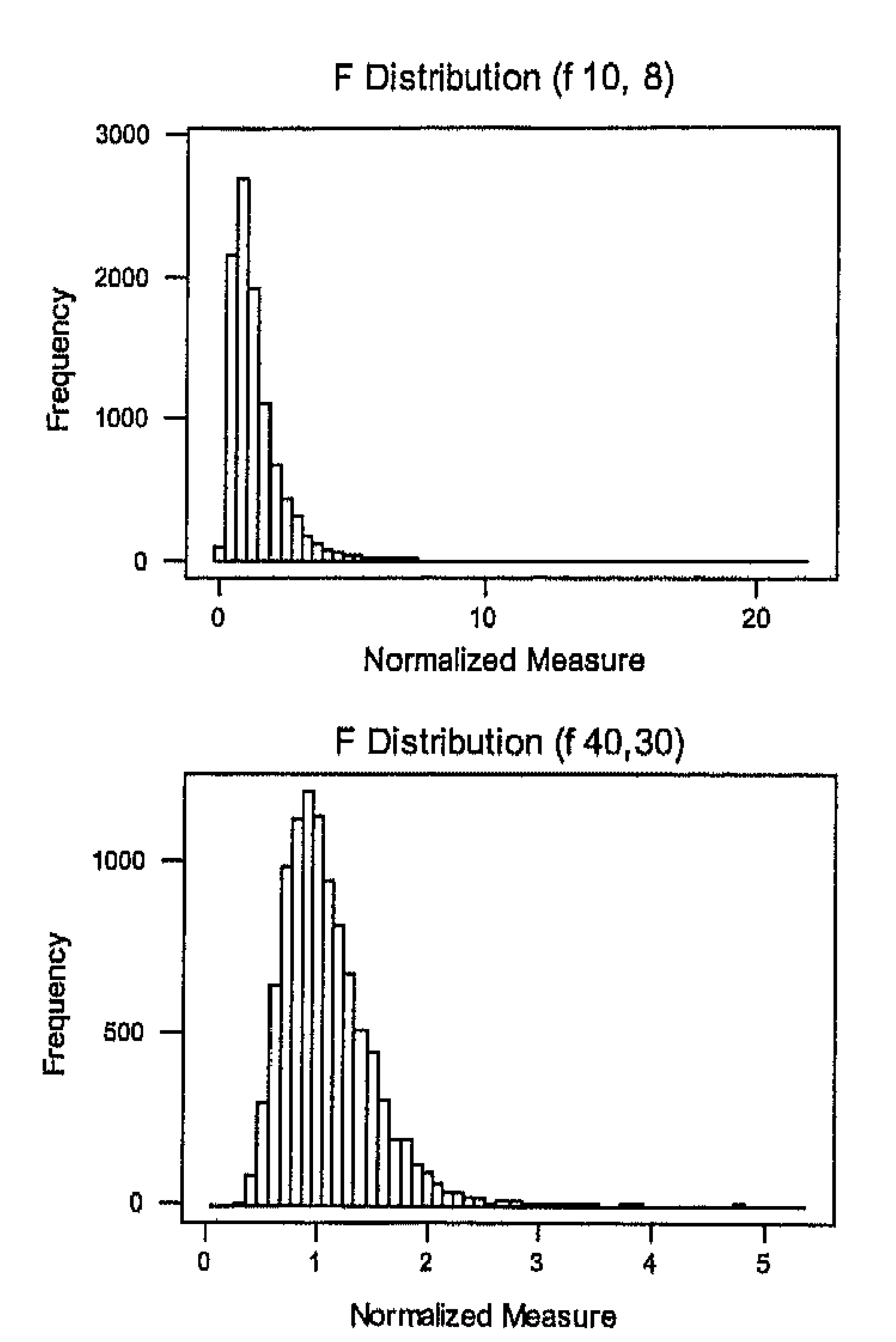
لرفض أو قبول فرضية العدم نحسب الإحصائية $F$ التالية :

$$F = \frac{S_1^2}{S_2^2}$$

حيث $s_1^2$ و $s_2^2$ هي تقديرات تباينات العينات لمجتمعين. إن نسبة التباين لمجتمعين طبيعيين هي أيضاً متغير عشوائي يتبع التوزيع $F$. يوضح الشكل (٥,٧) التوزيع $F$. وكما هو الحال مع التوزيع $t$ فإن التوزيع $F$ يختلف ببارامترين مثل أحجام العينات لمجتمعين. يبين الجدول (٥,١) جزءًا من الجدول $F$ حيث هناك حاجة إلى درجتي حرية $dn$ و $dd$ لتحديد إحدى قيم $F$ في الجدول $F$. يتم في هذا الجدول إعطاء إدخالات

الجدول لمستويات أهمية (قيم $\alpha$ ) تساوي 0.05 و 0.01. ويتم إعطاء قيم $F$ المرتبطة بمستويات الأهمية 0.05 و 0.01 بخط نظامي وبخط سميك مائل على التوالي. وهكذا بالنسبة لأي درجتي حرية هناك قيمتان لـ $F$ واحدة لمستوى الثقة 95% وواحدة لمستوى الثقة 99%.

لاستخدام الجدول $F$ مع الاختبار $F$ نقوم بتقدير إحصائية $F$ باستخدام تقديرات تباين العينة من كل مجتمع من المجتمعين اللذين نحاول مقارنتهما. لاحظ أنه من أجل استخدام هذا الجدول $F$ ينبغي وضع التباين الأكبر من التباينين في البسط من المعادلة المذكورة سابقاً.




**الشكل (٥,٧). المدرج الإحصائي لعينات مأخوذة من توزيعين F مختلفين. في اللوحة العليا تكون درجتا الحرية 10 و 8. في اللوحة السفلى فإن درجتي الحرية هما 40 و 30.**

الجدول (١ ، ٥). قيم من التوزيع $F$ لمساحات $\alpha$ في الذيل إلى اليمين من $F(dn, dd, \alpha)$ .

| dd | dn: 1 | 2 | 3 | 4 | 5 | ... | 10 | 11 |
|---|---|---|---|---|---|---|---|---|
| **dd** | **1** | **2** | **3** | **4** | **5** | **...** | **10** | **11** |
| **1** | 161, **4052** | 200, **4999** | 216, **5403** | 225, **5625** | 230 **5764** | | 242, **6056** | 243, **6082** |
| **2** | 18.5, **98.49** | 19.00, **99.01** | 19.16, **99.17** | 19.25, **99.25** | 19.30, **99.30** | | 19.39, **99.40** | 19.40, **99.41** |
| **3** | 10.13, **34.12** | 9.55, **30.81** | 9.28, **29.46** | 9.12, **28.71** | 9.01, **28.24** | | 8.78, **27.23** | 8.76, **27.13** |
| **4** | 7.71, **21.20** | 6.94, **18.00** | 6.59, **16.69** | 6.39, **15.98** | 6.26, **15.52** | | 5.96, **14.54** | 5.93, **14.45** |
| **5** | 6.61, **16.26** | 5.79, **13.27** | 5.41, **12.06** | 5.19, **11.39** | 5.05, **10.97** | | 4.74, **10.05** | 4.70, **9.96** |
| ... | | | | | | | | |
| **10** | 4.96, **10.04** | 4.10, **7.56** | 3.71, **6.55** | 3.48, **5.99** | 3.33, **5.64** | | 2.97, **4.85** | 2.94, **4.78** |
| **12** | 4.75, **9.33** | 3.88, **6.93** | 3.49, **5.95** | 3.26, **5.41** | 3.11, **5.06** | | 2.76, **4.30** | 2.72, **4.22** |
| **15** | 4.54, **8.68** | 3.68, **6.36** | 3.29, **5.42** | 3.06, **4.89** | 2.90, **4.56** | | 2.55, **3.80** | 2.51, **3.73** |
| ... | | | | | | | | |
| **20** | 4.35, **8.10** | 3.49, **5.85** | 3.10, **4.94** | 2.87, **4.43** | 2.71, **4.10** | | 2.35, **3.37** | 2.31, **3.30** |
| **50** | 4.03, **7.17** | 3.18, **5.06** | 2.79, **4.20** | 2.56, **3.72** | 2.40, **3.41** | | 2.02, **2.70** | 1.98, **2.62** |
| **100** | 3.94, **6.90** | 3.09, **4.82** | 2.70, **3.98** | 2.46, **3.51** | 2.30, **3.20** | | 1.92, **2.51** | 1.88, **2.34** |
| **200** | 3.89, **6.76** | 3.04, **4.71** | 2.65, **3.38** | 2.41, **3.41** | 2.26, **3.11** | | 1.87, **2.41** | 1.83, **2.34** |
| **∞** | 3.84, **6.64** | 2.99, **4.60** | 2.60, **3.78** | 2.37, **3.32** | 2.21, **3.02** | | 1.83, **2.32** | 1.79, **2.24** |

$dn$ = درجات الحرية للبسط ؛ $dd$ = درجات الحرية للمقام ؛ α = المساحة في ذيل التوزيع إلى يمين $F(dn, dd, \alpha) = 0.05$ أو $0.01$.

تابع الجدول (١ , ٥).

| dn | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **∞** | **200** | **100** | **50** | **40** | **30** | **20** | **...** | **14** | **12** |
| 254 | 254 | 253 | 252 | 251 | 250 | 248 | | 245 | 244 |
| **6366** | **6352** | **6334** | **6302** | **6286** | **6258** | **6208** | | **6142** | **6106** |
| 19.50 | 19.49 | 19.49 | 19.47 | 19.47 | 19.46 | 19.44 | | 19.42 | 19.41 |
| **99.50** | **99.49** | **99.49** | **99.48** | **99.48** | **99.47** | **99.45** | | **99.43** | **99.42** |
| 8.53 | 8.54 | 8.56 | 8.58 | 8.60 | 8.62 | 8.66 | | 8.71 | 8.74 |
| **26.18** | **26.18** | **26.23** | **26.30** | **26.41** | **26.50** | **26.69** | | **26.92** | **27.05** |
| 5.63 | 5.65 | 5.66 | 5.70 | 5.71 | 5.74 | 5.80 | | 5.87 | 5.91 |
| **13.46** | **13.52** | **13.57** | **13.69** | **13.74** | **13.83** | **14.02** | | **14.24** | **14.37** |
| 4.36 | 4.38 | 4.40 | 4.44 | 4.46 | 4.50 | 4.56 | | 4.64 | 4.68 |
| **9.02** | **9.07** | **9.13** | **9.24** | **9.29** | **9.38** | **9.55** | | **9.77** | **9.89** |
| | | | | | | | | | |
| 2.54 | 2.56 | 2.59 | 2.64 | 2.67 | 2.70 | 2.77 | | 2.86 | 2.91 |
| **3.91** | **3.96** | **4.01** | **4.12** | **4.17** | **4.25** | **4.41** | | **4.60** | **4.71** |
| 2.30 | 2.32 | 2.35 | 2.40 | 2.42 | 2.46 | 2.54 | | 2.64 | 2.69 |
| **3.36** | **3.41** | **3.46** | **3.56** | **3.61** | **3.70** | **3.86** | | **4.05** | **4.16** |
| 2.07 | 2.10 | 2.12 | 2.18 | 2.21 | 2.25 | 2.33 | | 2.43 | 2.48 |
| **2.87** | **2.92** | **2.97** | **3.07** | **3.12** | **3.20** | **3.36** | | **3.56** | **3.67** |
| | | | | | | | | | |
| 1.84 | 1.87 | 1.90 | 1.96 | 1.99 | 2.04 | 2.12 | | 2.23 | 2.28 |
| **2.42** | **2.47** | **2.53** | **2.63** | **2.69** | **2.77** | **2.94** | | **3.13** | **3.23** |
| 1.44 | 1.48 | 1.52 | 1.60 | 1.63 | 1.69 | 1.78 | | 1.90 | 1.95 |
| **1.68** | **1.76** | **1.82** | **1.94** | **2.00** | **2.10** | **2.26** | | **2.46** | **2.56** |
| 1.28 | 1.34 | 1.39 | 1.48 | 1.51 | 1.57 | 1.68 | | 1.79 | 1.85 |
| **1.43** | **1.51** | **1.59** | **1.73** | **1.79** | **1.89** | **2.06** | | **2.26** | **2.36** |
| 1.19 | 1.26 | 1.32 | 1.42 | 1.45 | 1.52 | 1.62 | | 1.74 | 1.80 |
| **1.28** | **1.39** | **1.48** | **1.62** | **1.69** | **1.79** | **1.97** | | **1.17** | **2.28** |
| 1.00 | 1.17 | 1.24 | 1.35 | 1.40 | 1.46 | 1.57 | | 1.69 | 1.75 |
| **1.00** | **1.25** | **1.36** | **1.52** | **1.59** | **1.69** | **1.87** | | **2.07** | **2.18** |

نقوم الآن بمقارنة الإحصائية $F$ المُقدَّرة مع الإدخالات في الجدول $F$ المرتبطة بدرجات الحرية $dn$ و $dd$ ومستوى الثقة المناسب (يتم تقديم قيم $F$ الـ 95% و 99% فقط في الجدول المُقدَّم). $dn = n_1$ و $dd = n_2$ هي عدد العينات في كل مجتمع مع $n_1$ هو عدد العينات في المجتمع ذي التباين الموضوع في البسط.

إذا أردنا رفض فرضية العدم المبينة سابقاً يجب أن تكون الإحصائية $F$ المحسوبة أكبر من $F(\alpha, dn, dd)$ في الجدول لرفض $H_0$ بثقة مقدارها $(1-\alpha)\times 100\%$. درجات الحرية $dn = n_1$ هي القيمة المُستخدمة لتحديد موقع إدخال الجدول في الاتجاه الأفقي (البسط) و $dd = n_2$ هي درجات الحرية المُستخدمة لتحديد موقع إدخال الجدول في الاتجاه العمودي (المقام).

**مثال (٢ , ٥)**

الاختبار $F$

| | المجتمع الإحصائي A $(N_1 = 9)$ | المجتمع الإحصائي B $(N_2 = 9)$ |
|---|---|---|
| المتوسط | 0.026 | 0.027 |
| التباين | $2.0\times 10^{-5}$ | $7.4\times 10^{-5}$ |

وبالأخذ في الاعتبار البيانات الواردة في الجدول أعلاه حيث جمعنا 9 عينات من كل مجتمع من المجتمعين A و B يمكننا تقدير الإحصائية $F$ بحيث تكون :

$$F = (7.4E-5)/(2.0E-5) = 3.7$$

وباستخدام الجدول $F$ الذي لديه إدخالات جدول لكل من مستويات الثقة 95% و 99% نجد أن إدخال الجدول بالنسبة إلى $\alpha$ يساوي 0.05 و $df$ 9 لكل من المجتمعين هي :

$$F(0.05, 9, 9) = 3.18$$

تتجاوز القيمة 3.7 المُقدَّرة إدخال الجدول 3.18 ؛ وبالتالي فقد نرفض $H_0$ بثقة 95% ونقبل أن المجتمع B له تباين أكبر بكثير من المجتمع A. ومع ذلك نلاحظ أن إدخال الجدول من أجل $F(0.01, 9, 9)$ يساوي 5.35 وهو أكبر من القيمة المُقدَّرة لدينا. وهكذا فإننا لا نستطيع رفض فرضية العدم بثقة مقدارها 99%.

## (٥,٣) مقارنة ثلاثة متوسطات أو أكثر للمجتمعات الإحصائية

**COMPARISON OF THREE OR MORE POPULATION MEANS**

ناقشنا حتى الآن التحليل الإحصائي لمقارنة المتوسطات والتباينات (أو الانحرافات المعيارية) لمجتمعين على أساس جمع مجتمعين من البيانات. قد يكون هذان المجتمعان من البيانات مستقلين عن بعضهما بعضاً أو قد يكونان تابعين لبعضهما البعض والتي نشير إليها بالبيانات المزدوجة أو المحجوبة. هناك طريقة أخرى للتفكير في هذا الازدواج وهي تسميتها حجب الوحدة التجريبية وهذا يعني أن كلا المجتمعين من البيانات تم جمعهما من نفس الوحدات التجريبية ولكن تحت ظروف مختلفة. على سبيل المثال قد تكون الوحدات التجريبية كائنات بشرية أو حيوانات أو أجهزة طبية أو خطوط تصنيع أو مستعمرات خلية أو دارات إلكترونية وأكثر من ذلك. إذا كانت الوحدة التجريبية على سبيل المثال كائنات بشرية فإنه يمكن أخذ مجموعتي البيانات قبل وبعد إعطاء الدواء أو قبل وبعد اشتراك الأشخاص في تمرين. وفي مثال آخر قد نقيس نوعين مختلفين من البيانات مثل معدل ضربات القلب وضغط الدم كل منهما من نفس مجتمع الأشخاص. ويُشار إلى هذا بالمقاييس المتكررة وذلك لأننا نأخذ عدة مجموعات من المقاييس من نفس الوحدات التجريبية. في المقاييس المتكررة فإن شيئاً ما في الظروف التجريبية أو الوحدة التجريبية أو القياس الذي يجري جمعه يتباين من إحدى مجموعات المقاييس إلى التي تليها ولكن يتم حجب أخذ العينات من قبل الوحدة التجريبية.

نحتاج في العديد من تطبيقات الهندسة الطبية الحيوية إلى مقارنة المتوسطات من ثلاثة أو أكثر من المجتمعات الإحصائية أو العمليات أو الشروط. في مثل هذه الحالات نستخدم طريقة تحليل إحصائي تُدعى تحليل التباين أو أنوفا (ANOVA). على الرغم من أن الاسم يعني أن المرء يقوم بتحليل التباينات فإن الاستنتاجات التي تنبع من هذا التحليل هي بخصوص الاختلاف الكبير في المتوسطات لثلاثة أو أكثر من المجتمعات الإحصائية أو العمليات.

**(١,٣,٥) التجارب أحادية العامل One-Factor Experiments**

نبدأ مناقشاتنا لأنوفا (ANOVA) بمناقشة تصميم وتحليل التجارب أحادية العامل. نأخذ العينات في مثل هذه التجارب من ثلاثة مجتمعات إحصائية أو أكثر يتغير فيها عامل واحد من مجتمع إلى آخر.

وكما ورد في الفصل الثاني ينبغي أن يشمل التصميم التجريبي العشوائية والحجب. تضمن العشوائية عدم إدخال انحياز إلى البيانات بسبب ترتيب الآثار في جمع البيانات. وعلاوة على ذلك يساعدنا الحجب على التقليل من تأثير التغير بين الأشخاص على الاختلافات بين المجتمعات الإحصائية

إن بعض تحديات في الهندسة الطبية الحيوية التي قد تتطلب استخدام أنوفا تشمل ما يلي :

١ - مقارنة زمن الاهْتِراء بين زرعات الورك لمواد مختلفة : في هذا المثال المعالجة = مادة الزرع (على سبيل المثال التيتانيوم والفولاذ وراتنجات البوليمر).

٢- مقارنة متواليات نبضة الـ MR في القدرة على تصوير تلف الأنسجة بعد السكتة الدماغية : في هذا المثال المعالجة = متوالية نبضة الـ MR.

٣- مقارنة قدرة العديد من الأدوية على خفض ضغط الدم المرتفع : في هذا المثال المعالجة = الدواء.

### (٥,٣,١,١) مثال لتحدٍّ في الهندسة الطبية الحيوية

**Example of Biomedical Engineering Challenge**

أحد الأمثلة على مشكلة الهندسة الطبية الحيوية التي قد تستخدم أنوفا لاختبار الفروض هو دراسة آليات الانعكاس (رد الفعل) في المرضى المصابين بالحبل الشوكي [14]. إن أحد اهتمامات الباحثين هو تبعية انثناء الورك (عزم) لحركة الكاحل في المرضى الذين يعانون من إصابات الحبل الشوكي. في هذا المثال هناك بالفعل مجتمعان من العوامل التجريبية التي قد تؤثر على ثني الورك:

١- مجال حركة الكاحل: في هذه الحالة المعالجة = مجال تمديد الكاحل.

٢- سرعة ثني الكاحل: في هذه الحالة المعالجة = سرعة تمديد الكاحل.

ربما نقيّم عاملا واحدًا في الزمن باستخدام أنوفا أحادية العامل أو أحادية الاتجاه. أو ربما نقيّم تأثير كل من العاملين على متوسط ثني الورك باستخدام أنوفا أحادية العامل أو أحادية الاتجاه. وفي كلتا الحالتين لدينا فرضيات العدم والفرضيات البديلة لتأثير كل عامل مثل مجال حركة الكاحل على متوسط المجتمع الإحصائي مثل ثني الورك. إن فرضية العدم هي أنه لا يوجد اختلاف كبير في متوسط ثني الورك عبر حركة الكاحل.

$$H_0 : \mu_1 = \mu_2 = \mu_3 .... = \mu_n$$

وتنص الفروض البديلة $H_1$ على أن متوسطين على الأقل لمجتمعين مثل ثني الورك لحركتين مختلفتين للكاحل يختلفان كثيراً عن بعضهما البعض.

إن التصميم الأساسي للتجربة أحادية العامل مُعطى في الجدول (٥,٢). في هذه الحالة هناك $k$ معالجة في عامل واحد. $k$ هو عدد المعالجات $y_{ij}$ هو العينة $j$ الفردية أو نقطة البيانات للمعالجة رقم $i$ $\overline{y}_i$ هو متوسط العينة للمعالجة رقم $i$ و $\sigma_{yi}^2$ هو تباين العينة للمعالجة رقم $i$ (مقتبس من [3]).

الجدول (٢ , ٥). تجربة أحادية العامل.

| المعالجة | الملاحظات | متوسط العينة | تغيُّر العينة |
|---|---|---|---|
| 1 | $y_{11}, y_{22}, \ldots, y_{1n}$ | $\bar{y}_1$ | $\sigma^2 y_1$ |
| 2 | $y_{21}, y_{22}, \ldots, y_{2m}$ | $\bar{y}_2$ | $\sigma^2 y_2$ |
| ... | | | |
| $I$ | $y_{i1}, y_{i2}, \ldots, y_p$ | $\bar{y}_k$ | $\sigma^2 y_i$ |
| ... | | | |
| $K$ | $y_{k1}, y_{k2}, \ldots, y_{kq}$ | $\bar{y}_k$ | $\sigma y_k$ |

على سبيل المثال قد نكون مهتمين بمقارنة فقدان الوزن لثلاثة أنواع مختلفة من حبوب الحمية. في هذه الحالة $k = 3$. تحت عمود فقدان الوزن سيجد المرء فقدان الوزن لأشخاص فرديين استخدموا واحدة من حبوب الحمية الثلاث. لاحظ في هذا المثال أن عدد العينات يختلف في كل نوع من أنواع الأدوية. وفي تجربة متوازنة ينبغي أن يكون هناك عدد متساو من العينات لكل نوع من أنواع المعالجات. لدينا في العمودين الثالث والرابع متوسط العينة والانحراف المعياري لفقدان الوزن لكل حبة من حبوب الحمية. والسؤال الذي نحاول الإجابة عليه هو ما إذا كان هناك اختلاف كبير في متوسط فقدان الوزن كتابع لنوع حبة الحمية.

| حبة الحمية | فقدان الوزن | المتوسط | الانحراف المعياري |
|---|---|---|---|
| بلاسيبو (Placebo) | 10, 12, 5, 8, 5, 20 | 10.0 | 5.62 |
| الدواء B | 30, 5, 12, 20 | 16.75 | 10.75 |
| الدواء C | 2, 10, 5, 5, 10, 20, 25, 40 | 14.63 | 12.92 |

نقوم بعدد من الافتراضات عند استخدام أنوفا لمقارنة الاختلافات في المتوسطات بين ثلاثة أو أكثر من المجتمعات الإحصائية أو العمليات:

١- يتم تحديد الأشخاص بشكل عشوائي بالنسبة لمعالجة محددة (في هذه الحالة حبوب الحمية).

٢- إن المجتمعات الإحصائية أو العمليات (مثل فقدان الوزن) مُوزَّعة بشكل طبيعي تقريباً.

٣- إن التباين متساوٍ تقريباً عبر مجموعات المعالجة.

لم يتم في هذا المثال المحدد لفقدان الوزن استخدام الحجب. وبعبارة أخرى لم يحصل نفس الشخص في أي حال على أكثر من معالجة واحدة أو حبة حمية واحدة. قد يُحدِث هذا في نهاية الأمر مشكلة ويؤدي إلى استنتاجات خاطئة لأننا لا نأخذ في الاعتبار التغيّر بين الأشخاص كاستجابة لأدوية الحمية. على سبيل المثال قد نفترض أن مقدار فقدان الوزن مرتبط ببداية الوزن. من الممكن أن يكون لجميع الأشخاص المحددين للدواء A بداية أوزان أقل من الأشخاص الذين يتم إعطاؤهم الدواء B. وهكذا قد تكون الاختلافات في فقدان الوزن كبيرة ولكن قد يكون لها علاقة بسيطة بحبوب الحمية الفعلية. وقد يكون الاختلاف في الوزن بسبب الاختلافات في بداية الوزن بدلاً من حبوب الحمية. ونظراً لأن التجربة السابقة لا تستخدم الحجب فإن التغيّر بين الأشخاص الذي لا يتم أخذه في الاعتبار قد يُريك استنتاجاتنا بخصوص فعالية حبوب الحمية.

والسؤال الذي نحاول معالجته باستخدام أنوفا هو "هل متوسطات المعالجات $k$ متساوية أو هل تم أخذ العينات من معالجات (مجتمعات إحصائية) ذات متوسطات مختلفة؟"

نفترض داخل كل معالجة $i$ (أو مجتمع $i$) أن التغيّر عبر العينات التي تم ملاحظتها $Y_{ij}$ متأثر بمتوسط المجتمع (المعالجة) $\mu_i$ والمتغيرات العشوائية المستقلة $e_{ij}$ المُوزَّعة بشكل

طبيعي مع متوسط يساوي صفر وتباين $\sigma^2$. وبعبارة أخرى يمكن التعبير عن العينات $Y_{ij}$ التي تم جمعها لكل تجربة $j$ داخل كل معالجة كما يلي :

$$Y_{ij} = \mu_i + e_{ij}$$

نحاول بواسطة نموذجنا تحديد مقدار التغيُّر في $Y$ بسبب العامل أو المعالجة (مجتمع بمتوسط $\mu_i$) وكم هو هذا المقدار بسبب التأثيرات العشوائية $e_{ij}$ التي لا نستطيع التحكم بها أو التي لم يتم التقاطها في النموذج الذي تم تقديمه أعلاه.

عندما نجري أنوفا لتجربة أحادية العامل يمكننا تنظيم التحليل والنتائج في الجدول التالي (الجدول ٥,٣):

**الجدول (٣ , ٥). أنوفا (ANOVA) أحادية العامل بـ $k$ معالجة (من دون حجب).**

| المصدر | $df$ | $SS$ | $MS$ | $F$ |
|---|---|---|---|---|
| المعالجة | $k-1$ | | | $MS_{treat}/MS_{error}$ |
| الخطأ | $N-k$ | | | |

$N$ = العدد الكلي للعينات في جميع المعالجات ؛ $k$ = عدد المعالجات ضمن عامل واحد ؛ $F$ = الإحصاء الذي سوف نقارنه مع الجداول $F$ (التوزيع $F$) إما لرفض أو قبول فرضية العدم ؛ $SS_{treatment}$ = مجموع المربعات بين المعالجات هو مقياس التغير فيما بين متوسطات المعالجات ؛ $SS_{error}$ = مجموع المربعات ضمن المعالجات هو مقياس لمجموع التباينات في جميع المعالجات $k$ ؛ $MS_{treatment} = SS_{treatment}/k-1$ ؛ $MS_{error} = SS_{error}/N-k$ .

ومن أجل $k$ معالجة فإن المعادلات المُحددة للعناصر SS هي التالية :

$$SS_{treatment} = \sum_{i=1}^{k} n_i \left(\overline{y}_i - \overline{y}_{grand}\right)^2$$

و

$$SS_{error} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Y_{ij} - \overline{Y}_i\right)^2 = \sum_{i=1}^{k} \sigma_{yi}^2$$

$\overline{Y}_{grand}$ هو متوسط العينة لجميع العينات في جميع العلاجات مجتمعة.

وأخيراً فإن الإحصائية $F$ الأكثر اهتماماً بالنسبة لنا في تقدير رفض أو قبول فرضية العدم هي:

$$F = MS_{treatment} / MS_{error}$$

لاختبار الفروض بمستوى أهمية مقداره $\alpha$ نقارن الآن الإحصائية $F$ المُقدَّرة مع التوزيع $F$ المتوفر في جداول $F$ لإدخال الجدول $F\,(\alpha; k-1, N-k)$. إذا كانت قيمة $F$ المُقدَّرة أكبر من إدخال الجدول فقد نرفض فرضية العدم بمستوى ثقة مقداره $(1-\alpha)\times 100\%$.

**مثال (٣ , ٥)**

لدى صانع صمامات القلب ثلاث عمليات مختلفة لإنتاج صمام الوريقة. تم اختيار عينات عشوائية من 50 صمامًا خمس مرات من كل نوع لعملية التصنيع. تم اختبار عيوب آليات الفتح لكل صمام. يتم تلخيص عدد الصمامات التالفة في كل عينة من 50 صمامًا بالجدول التالي:

| العملية A | العملية B | العملية C |
|---|---|---|
| 1 | 5 | 3 |
| 4 | 8 | 1 |
| 3 | 6 | 1 |
| 7 | 9 | 4 |
| 5 | 10 | 0 |

باستخدام $\alpha = 0.05$ نريد تحديد ما إذا كان متوسط عدد العيوب يختلف بين العمليات (المعالجة). (فرضية العدم $H_0$ هي أن متوسط عدد العيوب هو نفسه في جميع العمليات).

وللإجابة على السؤال نحتاج إلى إكمال جدول أنوفا التالي :

| **أنوفا (ANOVA) أحادية الاتجاه: العمليات A و B و C.** | | | | |
|---|---|---|---|---|
| **المصدر** | *df* | *SS* | *MS* | *f* |
| العامل | 2 | | | |
| الخطأ | 12 | | | |
| الإجمالي | 14 | | | |

نجد أولاً $SS_{factor}$ (= $SS_{treatment}$)

نحن نعلم أن

$$SS_{treatment} = \sum_{i=1}^{k} n_i \left(\overline{y}_i - \overline{y}_{grand}\right)^2$$

نعلم من البيانات المُعطاة أن هناك ثلاثة معالجات A و B و C ؛ وبالتالي $k = 3$. وسوف نحدد $A = 1$ و $B = 2$ و $C = 3$. ونعلم أيضاً أن $n_{\cdot 1} = n_{\cdot 2} = n_{\cdot 3} = 5$. يمكننا تقدير متوسط العينة لكل معالجة أو عملية للحصول على $\overline{y}_1 = 4, \overline{y}_2 = 7.6, \overline{y}_3 = 1.8$ ويمكننا أيضاً استخدام جميع العينات الخمسة عشر لإيجاد $\overline{y}_{grand} = 4.47$.

يمكننا الآن استخدام هذه التقديرات في المعادلة من أجل $SS_{treatment}$ :

$$SS_{treatment} = 5(4 - 4.47)^2 + 5(7.6 - 4.47)^2 + 5(1.8 - 4.47)^2 = 85.70$$

ونحل الآن من أجل $SS_{error}$ :

وبالأخذ في الاعتبار أن :

$$SS_{error} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Y_{ij} - \overline{Y}_i\right)^2$$

فإننا بحاجة إلى تقدير المجموع الداخلي لكل معالجة $i$ مكتوبة بواسطة المجموع الخارجي. وهكذا عندما $i = 1$

$$SS_1 = (1-4)^2 + (4-4)^2 + (3-4)^2 + (7-4)^2 + (5-4)^2 = 20;$$

وعندما $i = 2$

$$SS_2 = (5-7.6)^2 + (8-7.6)^2 + (6-7.6)^2 + (9-7.6)^2 + (10-7.6)^2 = 17.2;$$

وعندما $i = 3$

$$SS_3 = (3-1.8)^2 + (1-1.8)^2 + (1-1.8)^2 + (4-1.8)^2 + (0-1.8)^2 = 10.8 .$$

وبالجمع الآن لجميع المعالجات $k$ حيث $k = 3$ نجد :

$$SS_{error} = SS_1 + SS_2 + SS_3 = 48.0$$

إيجاد متوسط الأخطاء المربعة :

$$MS_{treatment} = SS_{treatment} / DF_{treatment} = 87.5/2 = 43.7$$

$$MS_{error} = SS_{error} / DF_{error} = 48.0/12.0 = 4$$

وأخيراً تُعطى الإحصائية $F$ بواسطة العلاقة التالية :

$$F = MS_{treatment} / MS_{error} = 43.7/4 = 10.92$$

ونكمل الآن جدول أنوفا.

| أنوفا (ANOVA) أحادية الاتجاه: العمليات A و B و C. | | | | |
|---|---|---|---|---|
| المصدر | $df$ | $SS$ | $MS$ | $f$ |
| العامل | 2 | 87.5 | 43.7 | 10.92 |
| الخطأ | 12 | 48.0 | 4.0 | |

نريد الآن تحديد ما إذا كان متوسط عدد العيوب يختلف فيما بين العمليات A و B و C. إن فرضية العدم هي أن متوسط عدد العيوب لا يختلف فيما بين العمليات. ولرفض هذه الفروض بمستوى ثقة مقداره 95% $\alpha = 0.05$ يجب أن تكون الإحصائية $F$

المُقدَّرة أكبر من $F(0.05,2,12)=3.83$ الموجودة في جداول التوزيع $F$. الإحصائية $F$ تساوي 10.92 وهي أكبر من 3.88 ؛ وبالتالي يمكننا رفض فرضية العدم ثقة مقدارها 95% ونخلص إلى أن عدد العيوب يختلف مع اختلاف عملية التصنيع وأن عينات المعالجة A و B و C تم أخذها من مجتمعات ذات متوسطات مختلفة. وفي الواقع يعرض جدول التوزيع $F$ قيمة من أجل $F(0.01,2,12)=6.93$. وهكذا يمكننا كذلك رفض فرضية العدم بمستوى ثقة مقداره 99%.

**مثال (٤ , ٥)**

يجري تقييم سرعة التقاط الصورة لأربعة أنواع من ماسحات الـ MR. ويلخص الجدول التالي السرعات التي تم قياسها (بالدقائق) لثلاث عينات لكل نوع من الماسحات.

| الماسح A | الماسح B | الماسح C | الماسح D |
|---|---|---|---|
| 2.0 | 4.0 | 3.0 | 6.0 |
| 1.8 | 4.5 | 2.5 | 5.5 |
| 2.7 | 5.5 | 2.0 | 3.5 |

هل يختلف متوسط سرعة التقاط الصورة فيما بين أنواع الماسح الأربعة؟ (استخدم $\alpha=0.01$).

للإجابة على هذا السؤال فقد نعمل من خلال العمليات الحسابية التي أجريناها في المثال (٥,٣). ومع ذلك كما اكتشف المرء في المثال السابق عندما يزداد حجم العينة فإن عدد الحسابات يزداد بسرعة. وهكذا فإن معظم الباحثين يستخدمون حزمة برامج إحصائية لتنفيذ حسابات ANOVA (أنوفا). على سبيل المثال استخدمنا حزمة البرامج

ميني تاب (Minitab Statistical Software, Release 13.32, Minitab, 2000) لتنفيذ ANOVA. تنتج أنوفا الجدول التالي :

| أنوفا (ANOVA) للسرعة | | | | | |
|---|---|---|---|---|---|
| المصدر | $df$ | $SS$ | $MS$ | $f$ | $\alpha$ |
| العامل | 3 | 19.083 | 6.361 | 9.07 | 0.006 |
| الخطأ | 8 | 5.613 | 0.702 | | |
| الإجمالي | 11 | 24.697 | | | |

ينتج عن الإحصائية $F$ لـ 9.07 مساحة مقدارها $\alpha = 0.006$ في الذيل الأيمن للتوزيع $F$. وبما أن $\alpha < 0.01$ وهي قيمة $\alpha$ التي كنا نختبر عندها الفروض، فإنه يمكننا رفض فرضية العدم وقبول هذه الفروض البديلة بأن واحداً على الأقل من الماسحات يختلف عن الماسحات المتبقية في متوسط سرعة الالتقاط.

لم نستخدم الحجب في التصميم التجريبي في الأمثلة المذكورة سابقاً. وبعبارة أخرى فإن المجتمعات التي جمعنا منها العينات اختلفت من معالجة إلى معالجة. في بعض التجارب مثل اختبار فقدان الوزن بسبب حبوب الحمية فإنه ليس عملياً أو ممكناً اختبار أكثر من نوع واحد من المعالجات على نفس الوحدة التجريبية.

ومع ذلك عندما يمكن استخدام الحجب ينبغي استخدامه لمقارنة المعالجات للتقليل من التغير بين الأشخاص أو الاختلافات (التي لا يمكن التحكم بها) في النتائج التجريبية. يبين الجدول (٥,٤) التصميم التجريبي لتجربة أحادية العامل تستخدم الحجب. تخضع جميع الوحدات التجريبية في هذا التصميم التجريبي للمعالجة في كل مرة. والنتيجة هي أنه لدينا الآن متوسطات معالجة ومتوسطات حجب. وهكذا عندما نقوم بتحليل أنوفا فقد نقوم باختبار مجتمعين مختلفين من الفرضيات. فرضية العدم

الأولى هي أنه لا يوجد اختلاف كبير في متوسطات المعالجة. وفرضية العدم الثانية هي أنه لا يوجد اختلاف كبير في المتوسطات عبر الأشخاص أو الوحدات التجريبية (التغيّر بين الأشخاص ليس كبيراً).

الجدول (٤ , ٥). عامل واحد بـ $k$ معالجة وحجب

| المعالجة (سرعة الكاحل) | رقم الشخص | | | | متوسط المعالجة |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| 10 | 5 | 2 | 5 | 6 | |
| 20 | 10 | 2 | 8 | 4 | |
| 30 | 15 | 8 | 15 | 10 | |
| متوسط الحجب | | | | | |

يمكن في هذا المثال أن تكون المعالجة سرعة الكاحل وقد يكون الحجب المريض بحيث يتم اختبار كل سرعة كاحل على كل شخص مريض. من المهم في مثل هذا التصميم جعل الترتيب الذي يتم بواسطته اختبار سرعات كاحل مختلفة عشوائياً بحيث لا يكون هناك انحياز في ثني الورك بسبب آثار ترتيب سرعة الكاحل أو اهتراء الجهاز.

لاحظ أن المثال سابقاً هو أيضاً تصميم تجريبي متوازن لأن هناك نفس الأعداد من نقاط البيانات في كل خلية من الجدول.

نفترض الآن داخل كل معالجة $i$ (أو مجتمع $i$) أن التغير عبر العينات التي تم ملاحظتها $Y_{ij}$ يتأثر بمتوسط المجتمع (المعالجة) $\mu_i$ وآثار الحجب $\beta_j$ والمتغيرات العشوائية المستقلة $e_{ij}$ المُوزَّعة بشكل طبيعي بمتوسط مقداره صفر وتباين مقداره $\sigma^2$. وبعبارة أخرى يمكن التعبير عن العينات $Y_{ij}$ التي تم جمعها لكل تجربة $j$ داخل كل معالجة $i$ كما يلي :

$$Y_{ij} = \mu_i + \beta_i + e_{ij}$$

وبعبارة أخرى نحاول تحديد مقدار التغير في $Y$ بسبب العامل أو المعالجة (مجتمع بمتوسط $\mu_i$) و كم هو بسبب آثار الحجب $\beta_j$ والآثار العشوائية $e_{ij}$ التي لا يمكننا التحكم بها أو لم يتم التقاطها في النموذج المُقدَّم سابقاً. نفترض في مثل هذا النموذج أن آثار المعالجة وآثار الحجب هي آثار مُضافة. وهذا ليس افتراضاً جيداً عندما تكون هناك آثار تفاعل بين المعالجة والحجب. تعني آثار التفاعل أن تأثير معالجة معينة قد يعتمد على حجب معين. عندما نعالج تجارب ثنائية العامل في المقطع التالي فسوف نناقش إلى حد كبير آثار التفاعل هذه.

يمكننا تلخيص أنوفا (ANOVA) لعامل واحد مع حجب بجدول أنوفا التالي (بافتراض عدم وجود أي آثار تفاعل) (الجدول ٥,٥) ؛ مقتبسة من [3]).

**الجدول رقم (٥ , ٥). أنوفا (ANOVA) أحادية العامل مع حجب.**

| المصدر | $df$ | $SS$ | $MS$ | $F$ |
|---|---|---|---|---|
| المعالجة | $k-1$ | | | $MS_{treat}/MS_{error}$ |
| الحجب | $b-1$ | | | $MS_{block}/MS_{error}$ |
| الخطأ | $(b-1)(k-1)$ | | | |

لاحظ أنه لدينا الآن قيمتين $F$ مُقدَّرتين للأخذ في الاعتبار. تتم مقارنة $F_{\text{treatment}}$ مع إدخال الجدول من أجل $F(\alpha; k-1, (b-1)(k-1))$ ويتم رفض فرضية العدم التي تنص على أنه ليس هناك فرق في متوسط ثني الورك عبر المعالجات إذا كانت قيمة $F$ المُقدَّرة أكبر من إدخال الجدول.

تتم مقارنة $F_{\text{block}}$ مع إدخال الجدول من أجل $F(\alpha; b-1, (b-1)(k-1))$ ويتم رفض فرضية العدم التي تنص على أنه ليس هناك فرق في متوسط ثني الورك عبر الأشخاص إذا كانت إحصائية $F$ المُقدَّرة أكبر من إدخال الجدول.

**مثال (٥ , ٥)**

تجربة أحادية العامل مع حجب.

تقوم شركة HeartSync بتصنيع أربعة أنواع من أجهزة إزالة الرجفان تختلف في قوة الصدمة الكهربائية التي يتم إعطاؤها لحادثة رجفان. تم تقسيم ما مجموعه 280000 مريض إلى أربع مجموعات كل منها 70000 مريض. وتم تخصيص كل مجتمع إلى أحد أجهزة إزالة الرجفان الأربعة وتم تسجيل عدد الصدمات التي فشلت في إزالة الرجفان لمدة أربع سنوات متتالية. وكانت النتائج على النحو التالي :

| سنة بعد الزراعة | الجهاز A | الجهاز B | الجهاز C | الجهاز D |
|---|---|---|---|---|
| 1 | 6 | 1 | 9 | 2 |
| 2 | 8 | 1 | 10 | 2 |
| 3 | 5 | 3 | 8 | 0 |
| 4 | 10 | 2 | 11 | 5 |

هناك نوعان من الأسئلة التي نتمنى معالجتها بهذه البيانات :

١ – باستخدام $\alpha = 0.01$ هل يختلف متوسط عدد الإخفاقات كثيراً كتابع لنوع الجهاز؟

٢ – باستخدام $\alpha = 0.01$ هل يختلف متوسط عدد الإخفاقات كثيراً كتابع لعدد السنوات بعد الزرع؟ وبعبارة أخرى هل عدد السنوات بعد الزرع مصدرٌ رئيسيٌّ للتغير بين المجتمعات؟

مرة أخرى بدلاً من تقدير الحسابات يدوياً قد نستخدم البرمجيات الإحصائية مثل Minitab للحصول على النتائج التالية باستخدام أنوفا مع حجب :

| أنوفا (ANOVA) لعدد الإخفاقات | | | | |
|---|---|---|---|---|
| المصدر | $df$ | $SS$ | $MS$ | $f$ |
| الجهاز | 3 | 173.188 | 57.729 | 35.67 |
| السنة | 3 | 20.688 | 6.896 | 4.26 |
| الخطأ | 9 | 14.563 | 1.618 | |
| الإجمالي | 15 | 208.438 | | |

إن الإحصائية $F$ للجهاز $F = 35.67$ أكبر من قيمة $F$ الحرجة $F(0.01,3,9)=6.99$ المُعطاة في جدول $F$. وهكذا يمكننا رفض فرضية العدم وقبول هذه الفروض البديلة بأن واحداً على الأقل من الأجهزة يختلف عن الأجهزة المتبقية في متوسط عدد الإخفاقات. الجزء الثاني من الأسئلة يختبر الفروض القائلة بأن متوسط عدد الإخفاقات يختلف اختلافاً كبيراً كدالة لعدد السنوات بعد الزرع (عامل الحجب). والإحصائية $F$ للسنة $F = 4.26$ هو أقل من قيمة $F$ الحرجة $F(0.01,3,9)=6.99$ المُعطاة في جدول $F$ وبالتالي فإننا نقبل فرضية العدم وهو ما يعني أن معدل الإخفاق لا يختلف بين عدد السنوات بعد الزرع.

**(٢ , ٣ , ٥) التجارب ثنائية العامل Two-Factor Experiments**

ناقشنا في التحدي في الهندسة الطبية الحيوية الذي يصف انعكاسات (ردود فعل) ثني الورك اثنين من العوامل التي تؤثر على ثني الورك : سرعة ومجال تمديد الكاحل.

في التجربة ثنائية العامل التي أدت إلى أنوفا ثنائية الاتجاه هناك عاملان يجري تغييرهما A و B حيث إن A لديه $a$ معالجات و B لديه $b$ معالجات وهناك $n$ عينة في كل تركيبة في A و B.

يُقال إن التجربة ثنائية العامل متقاطعة بشكل كامل إذا كانت هناك عينات تم جمعها لكل تركيبة من العوامل A و B. بالإضافة إلى ذلك يُقال إن التجربة متوازنة إذا كان لدينا نفس العدد من العينات لكل تركيبة من العوامل A و B

يوضح الجدول (٥,٦) أدناه تجربة ثنائية العوامل متوازنة ومتقاطعة بشكل كامل. يتم مقاطعة كل معالجة من المعالجتين ضمن العامل A مع كل معالجة من المعالجات الثلاثة ضمن العامل B. لاحظ أيضاً أنه لدينا ثلاث عينات لكل تركيبة من A و B.

الجدول رقم (٦ ,٥). تجربة ثنائية العامل

| العامل A | العامل B | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| 1 | 1.2, 1.4, 2.1 | 2.3, 2.2, 2.6 | 6.4, 5.8, 3.2 |
| 2 | 3.2, 4.1, 3.6 | 4.1, 4.3, 4.0 | 8.2, 7.8, 8.3 |

السؤال الذي نحاول معالجته بواسطة أنوفا ثنائية العامل هو ما إذا كانت هناك اختلافات في متوسطات المعالجات لكل عامل من العاملين. وبالإضافة إلى ذلك نريد أن نعرف ما إذا كانت هناك آثار تفاعل بحيث يكون هناك اختلافات كبيرة في المتوسطات كدالة لتفاعل الارتباط بين العوامل. وبعبارة أخرى هناك اختلافات كبيرة في متوسطات العينات وبالتالي في متوسطات المجتمعات الإحصائية عندما تحدث تركيبات معينة من العوامل A و B معاً.

وبمجرد جمع البيانات كما هو موضح في الجدول (٥,٦) أعلاه فقد نجري أنوفا ثنائية العامل لاختبار فرضيات العدم الثلاث التالية :

$$H_0 : \mu_{A1} = \mu_{A2} = \mu_{A3} = .... = \mu_{Aa};$$
$$H_0 : \mu_{B1} = \mu_{B2} = \mu_{B3} = .... = \mu_{Bb};$$
$$H_0 : \mu_{A1B1} = \mu_{A1B2} = \mu_{A1B3} = \mu_{A2B1} = \mu_{A2B2} = ... = \mu_{AaBb}$$

إن الفروض البديلة المرافقة لكل فرضية من فرضيات العدم الثلاث هي أن هناك فرقاً كبيراً في اثنين من متوسطات المجتمعات الإحصائية على الأقل لعامل معين أو تركيبة من العوامل.

يمكن تنظيم تحليل ونتائج أنوفا ثنائية العامل كما هو وارد في الجدول (٥,٧) [٣].

**الجدول (٧ , ٥). جدول أنوفا (ANOVA) ثنائية العامل (كل منها مع معالجات متعددة).**

| المصدر | $df$ | $SS$ | $MS$ | $F$ |
|---|---|---|---|---|
| A | $a-1$ | | | $MS_A / MS_{error}$ |
| B | $b-1$ | | | $MS_B / MS_{error}$ |
| AB | $(a-1)(b-1)$ | | | $MS_{AB} / MS_{error}$ |
| الخطأ | $ab(n-1)$ | | | |

إن معادلات $SS$ و $MS$ لكل عامل من العوامل وعوامل التفاعل هي خارج نطاق هذا الكتاب ولكن يمكن إيجادها في [3]. يستخدم الباحثون في الواقع حزمة برامج إحصائية شائعة مثل Minitab أو SPSS أو SAS لتقدير قيم $SS$ و $MS$ (بسبب العبء الحاسوبي) والإشارة ببساطة إلى الإحصائيات $F$ لرفض أو قبول فرضيات العدم.

نلاحظ أن هناك ثلاثة إحصائيات $F$ لاختبار كل فرضية من فرضيات العدم الثلاث التي سبق وصفها. ففي كل حالة نقارن إحصائية $F$ المُقدَّرة مع القيم في الجدول $F$ الذي يمثل التوزيع $F$. وبشكل أكثر تحديداً فإننا نقارن تقديرات قيم $F$ مع إدخالات الجدول التالي:

لاختبار $H_0 : \mu_{A1} = \mu_{A2} = \mu_{A3} = .... = \mu_{Aa}$ قارن $F_A$ مع $F(\alpha; a-1, ab(-1))$ ؛

لاختبار $H_0 : \mu_{B1} = \mu_{B2} = \mu_{B3} = .... = \mu_{Bb}$ قارن $F_B$ مع $F(\alpha; b-1, ab(-1))$ ؛

لاختبار $H_0 : \mu_{A1B1} = \mu_{A1B2} = \mu_{A1B3} = \mu_{A2B1} = \mu_{A2B2} = ... = \mu_{AaBb}$ قارن $F_{AB}$ مع $F(\alpha; (a-1)(b-1), ab(-1))$.

في كل اختبار من الاختبارات الثلاثة إذا كانت الإحصائية $F$ المُقدَّرة أكبر من إدخالات الجدول للتوزيع $F$ يمكن للمرء رفض فرضية العدم وقبول الفروض البديلة بثقة مقدارها $(1-\alpha)\times 100\%$.

**مثال (٦ , ٥)**

مثال على تجربة ثنائية العوامل سيتم تقييمها باستخدام أنوفا ثنائية العامل تحدث عندما يبحث المهندسون الطبيون الحيويون في فعالية علاج إعادة التأهيل والعلاج الدوائي في شفاء حركة طرف بعد السكتة الدماغية. تتضمن تفاصيل التصميم التجريبي ما يلي:

١ – العامل T: العلاج المُستخدم (هناك ثلاثة أنواع من العلاجات T1 و T2 و T3)؛

٢ – العامل D: الدواء المُستخدم (هناك ثلاثة أنواع من الأدوية D1 و D2 و D3)؛

٣ – يتم تحديد 36 مريضًا بشكل عشوائي إلى كل تركيبة من T و D؛

٤ – القياس: عدد الأيام لتحقيق معايير الشفاء.

| تصميم تجريبي لعاملَين | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| T3 | | | T2 | | | T1 | | |
| D3 | D2 | D1 | D3 | D2 | D1 | D3 | D2 | D1 |
| 7 | 15 | 9 | 16 | 8 | 22 | 13 | 25 | 20 |
| 10 | 10 | 12 | 19 | 10 | 16 | 12 | 16 | 15 |
| 9 | 9 | 8 | 11 | 9 | 17 | 22 | 10 | 18 |
| 9 | 10 | 8 | 21 | 11 | 12 | 10 | 20 | 24 |

الأسئلة التي نحاول معالجتها هي ما إذا كان هناك اختلاف كبير في متوسط أيام الشفاء لأنواع علاج إعادة التأهيل الثلاثة ومتوسط أيام الشفاء لأنواع العلاج الدوائي الثلاثة والاختلافات في متوسط أيام الشفاء للتركيبات التسعة من علاج إعادة التأهيل والعلاج الدوائي (تأثيرات التفاعل).

ينتج تحليل أنوفا الذي يتم إجراؤه باستخدام البرمجيات الإحصائية المعروفة باسم Minitab الجدول التالي الذي يلخص تحليل أنوفا. لاحظ أن هناك ثلاث إحصائيات $F$ مُقدَّرة. يمكننا استخدام قيم $F$ الثلاث لاختبار الفرضيات التالي:

$$H_0 : \mu_{T1} = \mu_{T2} = \mu_{T3};$$
$$H_0 : \mu_{D1} = \mu_{D2} = \mu_{D3};$$
$$H_0 : \mu_{T1D1} = \mu_{T1D2} = \mu_{T1D3} = \mu_{T2D1} = \mu_{T2D2} = ... = \mu_{T3D3}$$

**أنوفا (ANOVA) ثنائية الاتجاه لأيام الشفاء مقابل T و D.**

| المصدر | $df$ | $SS$ | $MS$ | $F$ | $\alpha$ |
|---|---|---|---|---|---|
| T | 2 | 337.4 | 168.7 | 11.4 | 0.000 |
| D | 2 | 36.2 | 18.1 | 1.23 | 0.309 |
| التفاعل | 4 | 167.8 | 41.9 | 2.84 | 0.043 |
| الخطأ | 27 | 398.3 | 14.8 | | |

لاحظ أن الإحصائية $F$ المُقدَّرة لعلاج إعادة التأهيل $F=11.44$ لا تتجاوز القيمة الحرجة عندما $\alpha < 0.05$ (في الواقع $\alpha$ هي أصغر من 0.0001)؛ وبالتالي فإننا نستنتج أن العينات لعلاجات إعادة التأهيل الثلاثة المختلفة تمثل مجتمعات إحصائية ذات متوسطات مختلفة. وبعبارة أخرى تؤدي علاجات إعادة التأهيل المختلفة إلى أيام شفاء مختلفة. ومع ذلك فإن الإحصائية $F$ المُقدَّرة للعلاج الدوائي لا تتجاوز قيمة $F$ الحرجة في الجدول عندما $\alpha < 0.05$ وبالتالي فإننا نقبل فرضية العدم بأن العلاج الدوائي لا يؤثر تأثيراً كبيراً على أيام الشفاء وأنه لم يتم أخذ العينات من مجموعات ذات متوسطات مختلفة. وعلاوة على ذلك فإن الإحصائية $F$ الثالثة $F = 2.84$ أكبر من قيمة $F$ في الجدول عندما $\alpha < 0.05$ مما يشير إلى أنه قد يكون هناك اختلاف كبير في أيام الشفاء بسبب التفاعل بين علاج إعادة التأهيل والعلاج الدوائي.

## (٣, ٣, ٥) إجراء توكي (Tukey) للمقارنات المتعددة

**Tukey's Multiple Comparison Procedure**

بعد أن أثبتنا أن هناك فارقاً كبيراً في المتوسطات عبر المعالجات ضمن عامل ما فقد نستخدم آخر الاختبارات المخصصة مثل اختبار HSD لتوكي للمعرفة المزدوجة والمقارنات المتعددة (Tukey's HSD multicomparison pair wise test) [3, 9]. تظهر أنوفا ببساطة أن هناك متوسط معالجة واحد على الأقل يختلف عن المتوسطات الأخرى. ومع ذلك فإن أنوفا لا تقدم معلومات محددة عن متوسط (متوسطات) المعالجة الذي يختلف عن متوسط (متوسطات) المعالجة الآخر. ويسمح اختبار HSD لتوكي بمقارنة الفروق الإحصائية في المتوسطات بين جميع أزواج المعالجات. وبالنسبة للتجربة أحادية العامل مع $k$ معالجة فإن هناك $k(k-1)/2$ مقارنة معرفة مزدوجة لاختبارها. والنقطة الهامة التي يجب ملاحظتها هي أنه إذا كانت $\alpha$ هي احتمال خطأ النوع I لمقارنة واحدة فإن احتمال القيام بخطأ النوع I على الأقل للمقارنات المتعددة أكبر بكثير. لذا إذا أردنا ثقة مقدارها $(1-\alpha)\times100\%$ لجميع مقارنات المعرفة المزدوجة الممكنة فإنه يجب أن نبدأ بـ $\alpha$ أصغر بكثير. يسمح إجراء توكي للمقارنات المتعددة بمثل هذه التعديلات في الأهمية عند إجراء مقارنات المعرفة المزدوجة.

# الفصل السادس

## الانحدار الخطي وتحليل الارتباط

## Linear Regression and Correlation Analysis

نهتم في بحوث الهندسة الطبية الحيوية أو التصميم، في كثير من الأحيان، فيما إذا كانت هناك ارتباط بين متغيرين، أو مجتمعين إحصائيين، أو عمليتين. ويمكن لهذه الارتباطات أن تعطينا معلومات حول العمليات البيولوجية الأساسية في الحالات العادية والمرضية وبالتالي تساعدنا على نمذجة العمليات، مما يسمح لنا بالتنبؤ بسلوك إحدى العمليات مع الأخذ في الاعتبار حالة عملية أخرى مرتبطة.

بالأخذ في الاعتبار مجموعتين من العينات، $X$ و $Y$، فإننا نطرح السؤال التالي: "هل المتغيران أو العمليتان العشوائيتان، $X$ و $Y$، مرتبطتان؟" وبعبارة أخرى، هل يمكن نمذجة $y$ كدالة خطية لـ $x$ بحيث يكون:

$$\hat{y}=mx+b$$

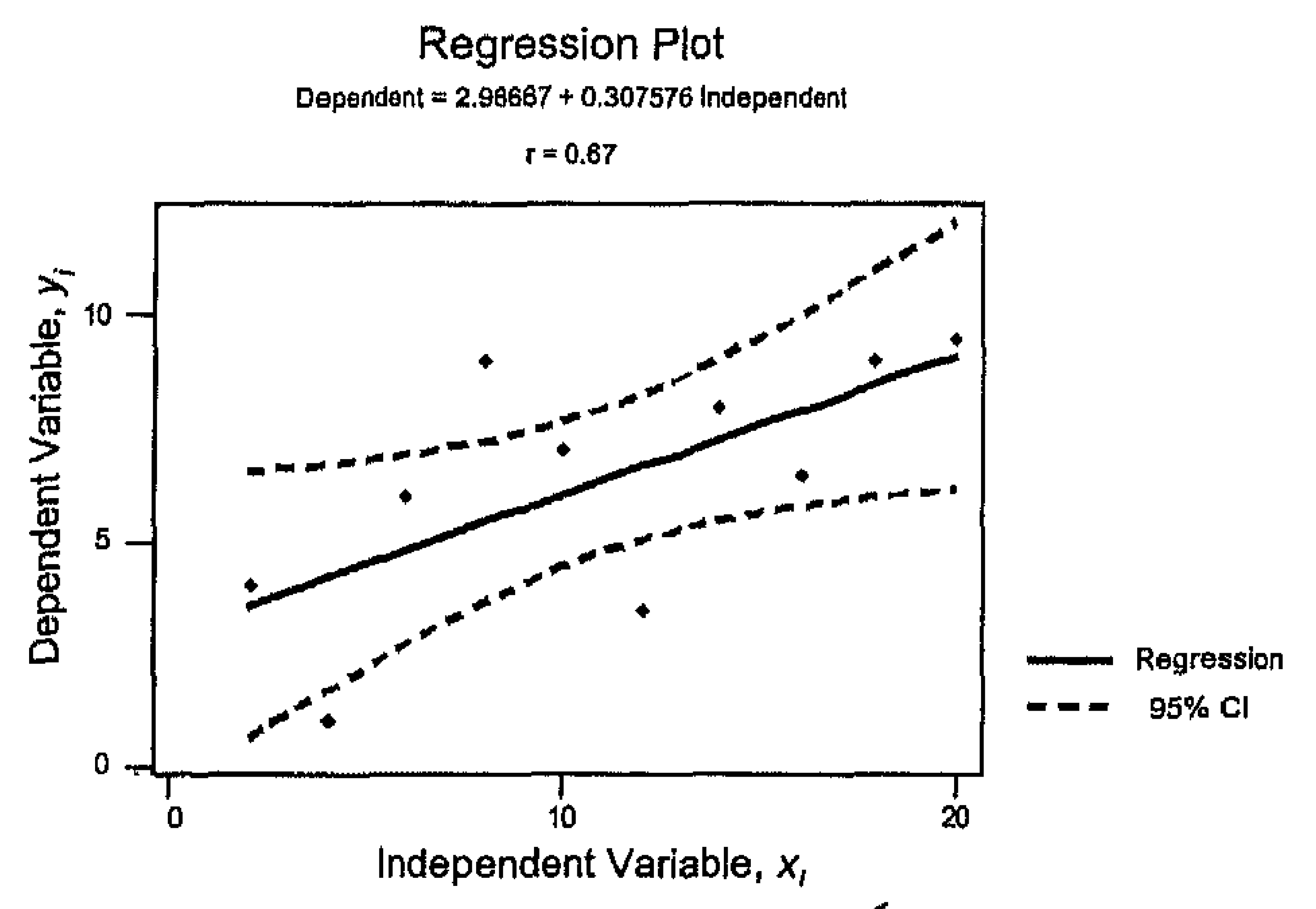
وبالنظر إلى المخطط التالي لبيانات تجريبية (الشكل ٦،١)، حيث تم رسم مجموعة البيانات، $y_i$، مقابل مجموعة البيانات، $x_i$:

نلاحظ أن البيانات تقترب من الوقوع على خط مستقيم. وهناك ميل إلى أن تصبح الاتجاه بحيث تزداد $y$ بما يتناسب مع الزيادات في $x$ وهدفنا هو تحديد الخط (النموذج الخطي) الذي يلائم هذه البيانات بأفضل ما يمكن ومقدار تقارب نقاط

البيانات المُقاسة بالنسبة إلى الخط الذي تم ملاءمته (الذي تم إنشاؤه بواسطة النموذج). وبعبارة أخرى، إذا كان الخط الذي تم نمذجته ملائمًا جدًّا للبيانات، فإننا نثبت أن y يمكن نمذجتها بدقة كدالة خطية لـ $x$، وبالتالي، يمكننا التنبؤ بـ $y$ مع الأخذ في الاعتبار $x$ باستخدام النموذج الخطي.

إن المدخل لملاءمة خط يتنبأ بالعملية $y$ بأفضل ما يمكن من العملية $x$، هو إيجاد البارامترات $m$ و $b$، التي تقلّل إلى الحد الأدنى من الخطأ بين بيانات النموذج والبيانات الفعلية بطريقة المربعات الصغرى:

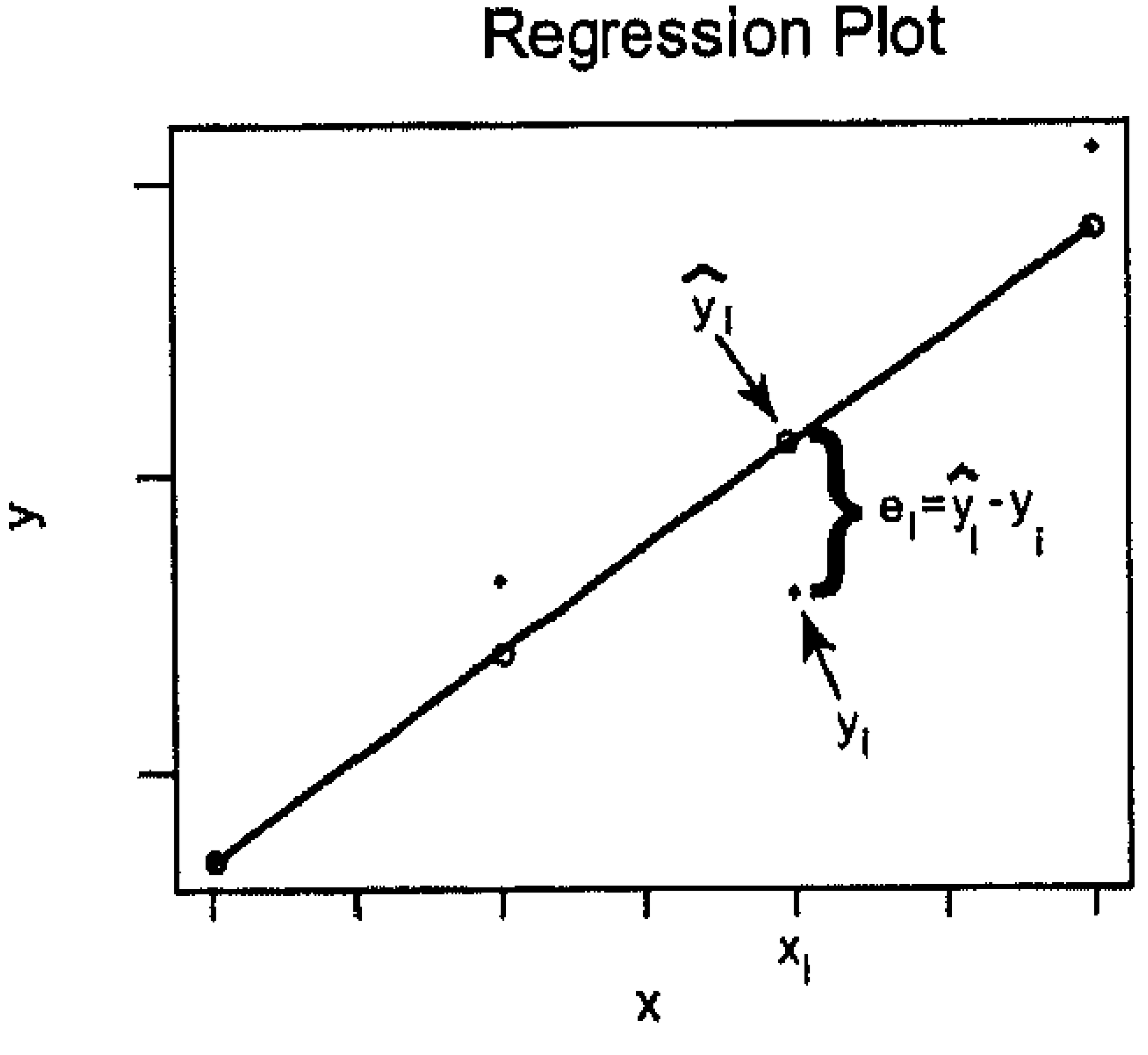
$$\min\left[(y-\hat{y})^2\right]$$




**الشكل (٦,١). نتائج خط الانحدار مُطبَّقة على العينات الموضحة بالمخطط المبعثر (النقاط السوداء). يوضح الخط الأسود المتصل خط الملاءمة الأفضل (بارامترات النموذج المذكورة فوق المخطط) على النحو الذي حدده خط الانحدار توضح المنحنيات الحمراء المنقطة فترة الثقة للميل (slope). وأخيراً، فإن القيمة $r$ هي معامل الارتباط.**

وبعبارة أخرى، كما هو موضح في الشكل (٦,٢)، فإن لكل قيمة مُقاسة للمتغير المستقل، $x$، سوف يكون هناك قيمة مُقاسة للمتغير التابع، $y$، بالإضافة إلى

القيمة المتوقعة أو التي تم نمذجتها لـ $y$، يُرمز لها بـ $\hat{y}$ ، التي سوف يحصل عليها المرء إذا تم استخدام المعادلة $\hat{y} = mx + b$ للتنبؤ بـ $y$. تدل المعادلة $e_i = y_i - \hat{y}_i$ على الأخطاء التي تحدث في كل زوج من $(x_i, y_i)$ عندما لا تكون القيمة التي تم نمذجتها مطابقة تماماً للقيمة المُتوقعة بسبب عوامل لم يتم أخذها في الاعتبار في النموذج (الضجيج، والآثار العشوائية، وعدم الخطية).




**الشكل (٦,٢). يمكن استخدام خط الانحدار لتقدير الخط المستقيم الذي "يلائم" على أفضل وجه نقاط البيانات المُقاسة (الدوائر المملوءة). تمثل $x_i$ و$y_i$ في هذا التوضيح المتغيرات المُقاسة المستقلة وغير المستقلة (التابعة)، على التوالي. يتم استخدام خط الانحدار لنمذجة المتغير التابع، y، كدالة خطية للمتغير المستقل، x. إن الخط المستقيم الذي يمر عبر نقاط البيانات المُقاسة هو نتيجة الارتداد الخطي حيث يتم تقليل الخطأ، $e_i$ ، إلى الحد الأدنى على كامل نقاط البيانات بين القيمة المتوقعة (الدوائر المفتوحة) للمتغير التابع، $\hat{y}_i$ ، والقيمة المُقاسة للمتغير التابع، $y_i$ .**

في محاولة لملاءمة الخط إلى البيانات التجريبية، فإن هدفنا هو تقليل هذه الأخطاء، $e_i$ ، إلى الحد الأدنى بين القيم المُقاسة والمُتوقعة للمتغير التابع، $y$. إن الطريقة المستخدمة في خط الانحدار والعديد من تقنيات النمذجة الطبية الحيوية الأخرى هي إيجاد بارامترات النموذج، مثل $m$ و $b$ ، التي تقلّل من مجموع الأخطاء المُربعة، $e_i^2$ ، إلى الحد الأدنى.

بالنسبة لخط الانحدار، فإننا نسعى إلى تقدير المربعات الصغرى لـ $m$ واستخدام التقريب التالي:

لنفترض أنه لدينا $N$ عينة لكل من العمليات $x$ و $y$. نحاول التنبؤ بـ $y$ من $x$ المُقاسة باستخدام النموذج التالي:

$$\hat{y} = mx + b$$

إن الخطأ في التنبؤ في كل نقطة بيانات، $x_i$، هو

$$error_i = y_i - \hat{y}_i$$

وفي طريقة المربعات الصغرى، فإننا نختار $m$ و $b$ للتقليل من مجموع الأخطاء المُربعة إلى الحد الأدنى:

$$(y_i - \hat{y}_i)^2 \qquad \text{عندما } i=1 \text{ إلى } N$$

لإيجاد حل شكل متقارب لـ $m$ و $b$، نستطيع كتابة صيغة لمجموع الأخطاء المُربعة:

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2$$

ثم نستبدل $\hat{y}_i$ بـ $(mx_i + b)$، نموذجنا، وإجراء عمليات التربيع [3, 5].

يمكننا بعد ذلك أخذ مشتقات الصيغة أعلاه بالنسبة إلى $m$ ومن ثم مرة أخرى بالنسبة إلى $b$. إذا وضعنا الصيَغ المُشتقة إلى الصفر لإيجاد القيم الدنيا، سيكون لدينا معادلتان بمجهولين، $m$ و $b$، ويمكننا ببساطة استخدام الجبر للحل بالنسبة للبارامترات المجهولة، $m$ و $b$. وسوف نحصل على الصيَغ التالية لـ $m$ و $b$ بدلالة $x_i$ و $y_i$ المُقاسة:

$$m = \frac{\sum_{i=0}^{N-1} x_i y_i - \left(\sum_{i=0}^{N-1} x_i\right)\left(\sum_{i=0}^{N-1} y_i\right) \Big/ N}{\sum_{i=0}^{N-1} x_i^2 - \left(\sum_{i=0}^{N-1} x_i\right)^2 \Big/ N}$$

و

$$b = \bar{y} - m\bar{x}$$

حيث

$$\bar{y}=\frac{1}{N}\sum_{i=0}^{N-1}y_i \qquad \text{و} \qquad \bar{x}=\frac{1}{N}\sum_{i=0}^{N-1}x_i$$

وبالتالي ، بعد حصولنا على البيانات المُقاسة ، يمكننا ببساطة استخدام المعادلات لـ $m$ و $b$ لإيجاد الخط ، أو النموذج الخطي ، ذي الملاءمة الأفضل.

**معامل الارتباط The Correlation Coefficient**

من المهم إدراك أن خط الانحدار يلائم خطًّا ما إلى أي مجموعتين من البيانات بغض النظر عن مدى جودة نمذجة البيانات بواسطة النموذج الخطي. وحتى لو كانت البيانات ، عندما يتم رسمها كمخطط مبعثر ، لا تبدو أنها تشبه الخط بشيء ، فإن خط الانحدار سوف يلائم الخط للبيانات. وكمهندسين طبيين حيويين ، علينا أن نسأل : "ما مدى جودة "ملاءمة" البيانات المُقاسة للخط المُقدَّر من خلال خط الانحدار ؟"

إن أحد مقاييس مدى جودة ملاءمة البيانات التجريبية للنموذج الخطي هو معامل الارتباط. يأخذ معامل الارتباط ، $r$، قيمة بين $-1$ و $1$ ويشير إلى مدى جودة ملاءمة النموذج الخطي للبيانات.

يمكن تقدير معامل الارتباط ، $r$، من البيانات التجريبية $x_i$ و $y_i$ ، وذلك باستخدام المعادلة التالية :

$$r=\frac{\sum_{i=0}^{N-1}(x_i-\bar{x})(y_i-\bar{y})}{\left[\sum_{i=0}^{N-1}(x_i-\bar{x})^2\sum_{i=0}^{N-1}(y_i-\bar{y})^2\right]^{1/2}}$$

حيث

$$\bar{y}=\frac{1}{N}\sum_{i=0}^{N-1}y_i \qquad \text{و} \qquad \bar{x}=\frac{1}{N}\sum_{i=0}^{N-1}x_i$$

من المهم أن نلاحظ أن $r = 0$ لا يعني أن العمليتين ، $x$ و $y$ ، مستقلتان. إنه يشير فقط إلى أن أي تبعية بين $x$ و $y$ ليست موصوفة أو مُنمذجة جيداً بواسطة علاقة خطية. ويمكن أن تكون هناك علاقة غير خطية بين $x$ و $y$. إن $r = 0$ يعني ببساطة أن $x$ و $y$ غير مترابطتين بالمعنى الخطي. وهذا يعني أن المرء قد لا يتوقع $y$ من $x$ باستخدام النموذج الخطي ، $y = mx + b$ .

إن أحد المقاييس المتصلة بمعامل الارتباط ، $r$ ، هو معامل التحديد ، $R^2$ ، الذي هو ملخص الإحصائية التي تخبرنا عن مدى جودة ملاءمة نموذج انحدار للبيانات. ويمكن استخدام $R^2$ كمقياس لجودة ملاءمة أي نموذج ارتداد ، وليس فقط خط الانحدار. وبالنسبة لخط الانحدار ، فإن $R^2$ هو مربع معامل الارتباط ، وله قيمة بين 0 و 1. يخبرنا معامل التحديد عن مقدار التغير في البيانات الذي يمكن تفسيره بواسطة بارامترات النموذج كجزء من مجموع التغير في البيانات.

ومن المهم معرفة أن الميل المُقدَّر للملاءمة الأفضل ومعامل الارتباط هو الإحصائيات التي قد تكون أو لا تكون مهمة. وبالتالي ، قد يتم إجراء الاختبارات $t$ لاختبار ما إذا كان الميل المُقدَّر من خلال ملاءمة خطية يختلف كثيراً عن الصفر [3]. وبالمثل ، قد يتم إجراء الاختبارات $t$ لاختبار ما إذا كان معامل الارتباط يختلف كثيراً عن الصفر. وأخيراً ، فقد نحسب أيضاً فترات الثقة للميل المُقدَّر [3].

# الفصل السابع

## تحليل القوى وحجم العينة
## Power Analysis and Sample Size

لقد ناقشنا حتى هذه اللحظة جوانب هامة من التصميم التجريبي، وملخص البيانات، والتحليل الإحصائي الذي يسمح لنا باختبار الفروض واستخلاص النتائج بمستوى معين من الثقة.

ومع ذلك، فإننا لم نعالج بعد مسألة في غاية الأهمية. والسؤال الذي نطرحه الآن هو: "كم ينبغي أن تكون العينة كبيرة لالتقاط التغير في المجتمع الإحصائي حتى تكون معدلات خطأ النوع I والنوع II صغيرة؟" وبعبارة أخرى، ما هو كبر العينة المطلوب بحيث يكون احتمال القيام بخطأ النوع I أو النوع II في رفض أو قبول فرضية العدم مقبولاً في ظل هذه الظروف. ووجود الحالات المختلفة يستدعي معدلات خطأ مختلفة. إن خطأ مثل تشخيص التهاب الحلق الناجم عن العقديات عندما تكون البكتيريا العقدية غير موجودة هو على الأرجح ليس خطيراً كما هو عند عدم تشخيص السرطان. والطريقة الأخرى لصياغة السؤال هي: "ما مقدار قوة التحليل الإحصائي في قبول أو رفض فرضية العدم؟"

إذا كان حجم العينة صغيراً جداً، فقد تكون النتيجة فقدان أثر ما (خطأ النوع II) لأننا لا نملك ما يكفي من القدرة في الاختبار لإثبات الأثر بثقة.

ومع ذلك من السهل جداً، عند اختيار حجم العينة، القول ببساطة إن حجم العينة ينبغي أن يكون كبيراً قدر الإمكان. وحتى لو كان للباحث إمكانية الحصول على عينات كثيرة بالقدر الذي يرغب فيه، فإن هناك اعتبارات وقيود عملية تحد من حجم العينة. إذا كان حجم العينة كبيراً جداً، فإن هناك مشاكل اقتصادية وأخلاقية يجب أخذها في الاعتبار. أولاً، هناك النفقات المرتبطة بالقيام بتجربة، مثل التجربة الإكلينيكية. وهناك التكاليف المرتبطة بالأفراد الذين يقومون بهذه التجارب، والوحدات التجريبية (الحيوانات، ومستعمرات الخلية، والتعويض عن وقت الإنسان)، وربما الأدوية والإجراءات الطبية الأخرى التي يتم إعطاؤها، وغيرها. وهكذا، كلما كان عدد العينات أكبر، كلما كانت النفقات أكبر. وعادة ما يكون القيام بالتجارب الإكلينيكية مكلفاً للغاية.

إن الاعتبار الثاني للحد من حجم العينة هو الموضوع الأخلاقي. إن العديد من التجارب والاختبارات الطبية الحيوية ذات الصلة يشمل عناصر بشرية أو حيوانية. وقد تتعرض هذه العناصر للأدوية أو العلاجات التجريبية التي تنطوي على بعض المخاطر، وفي حالة الدراسات على الحيوانات، قد يتم التضحية بالحيوان في نهاية التجربة. خلاصة القول، نحن لا نرغب في استخدام العناصر البشرية أو الحيوانية لأي سبب وجيه، وخاصة إذا كنا لا نكسب شيئاً فيما يتعلق بقدرة التحليل الإحصائي من خلال زيادة حجم العينة.

ونذكر مرة أخرى أن هناك نوعين من الأخطاء المرتبطة بإجراء التحليل الإحصائي:

١- النوع I: رفض $H_0$ عندما تكون $H_0$ صحيحة (احتمال خطأ النوع $I = \alpha$)؛

٢- النوع II: قبول $H_0$ عندما تكون $H_0$ خاطئة؛ فقدان أثر (احتمال خطأ النوع $II = \beta$).

## (٧,١) اختبار القوة
## POWER OF A TEST

عندما نشير إلى قدرة تحليل إحصائي أو اختبار معين، فإننا نحدد كمية الاحتمال، لقيمة $\alpha$ معينة (على سبيل المثال، 0.05)، بحيث يكشف الاختبار الإحصائي (الاختبار $t$، أو معامل الارتباط، أو أنوفا (ANOVA)، إلخ) أثراً حقيقياً. على سبيل المثال، إذا كان هناك حقاً اختلاف بين المتوسطات لمجتمعين، فما هو احتمال اكتشافنا لذلك الاختلاف بواسطة الاختبار $t$؟ إذا كانت فرصة حدوث خطأ النوع II هي $\beta$، فإن احتمال اكتشاف الاختلاف الحقيقي عندئذ هو ببساطة $1-\beta$، والذي نشير إليه بقدرة الاختبار. وفيما يلي بعض النقاط العملية حول قدرة الاختبار الإحصائي:

١- قدرة الاختبار = $1-\beta$ ؛

٢- صفر < القدرة < 1

٣- بشكل عام، نرغب أن تكون القدرة $\leq 0.8$ في التطبيقات العملية.

عند محاولة إنشاء حجم عينة لتجربة ما، يجب علينا التحديد مسبقاً مقدار القدرة الذي نريده للتحليل الإحصائي، الذي، بدوره، يتحدد بواسطة مقدار الخطأين للنوع I والنوع II اللذين نحن على استعداد للمخاطرة بهما. يمكننا، بالنسبة لبعض تطبيقات الهندسة الطبية الحيوية، المخاطرة بخطأ أكبر من غيرها من التطبيقات. إن فقدان التشخيص قد لا يكون مشكلة في إحدى الحالات أو قد يؤدي إلى حالة تهدد الحياة. إذا فشل جهاز إزالة الرجفان القابل للزرع في كشف نَظم القلب وصدمه كهربائياً، مثل الرجفان البطيني، فقد يعني هذا فقدان الحياة. من ناحية أخرى، فإن صدم القلب عندما لا يكون هناك رجفان، وهو ما قد يحدث عندما يتحسس الجهاز لنشاط كهربائي متولد عن عضلات الهيكل العظمي القريبة، قد يؤدي إلى الألم، وربما يؤدي أيضاً إلى بدء نَظم قلب خطير.

غالباً ما يكون من الصعب تقليل كل من α و β في الوقت نفسه. عادة ما يتم تقليل واحدة على حساب الأخرى. على سبيل المثال، قد نريد الخطأ إلى حد ما من جانب جهاز إزالة الرجفان القابل للزرع فائق الكشف للرجفان بحيث لا يتم فقدان أي حدوث للرجفان. في حالة الكشف عن سرطان الثدي باستخدام التصوير الشعاعي للثدي، فإن الكشف الزائد والكشف المتدني يمكن أن يكونا مشكلة. والكشف المتدني قد يعني زيادة احتمالات الوفاة. ومع ذلك، يمكن للإيجابيات الخاطئة أن تؤدي إلى إزالة لا لزوم لها من الأنسجة السليمة. وأحد أكبر التحديات التي تواجه الباحثين في المجال الطبي الحيوي هو محاولة إيجاد الاختلافات الحقيقية بين المجتمعات الإحصائية التي من شأنها تحسين التشخيص والكشف عن المرض أو أداء الأجهزة الطبية.

## (٧,٢) اختبارات القوى لتحديد حجم العينة
## POWER TESTS TO DETERMINE SAMPLE SIZE

تُستخدم اختبارات القدرة لتحديد حجم العينة، والأخذ في الاعتبار التأثير الذي ينبغي الكشف عنه بالتحليل الإحصائي، ومعدلات خطأ النوع I والنوع II المسموح بها، وتغيُّر المجتمعات الإحصائية التي يتم أخذ عيناتها في التجارب.

ولإجراء اختبار القدرة، يمكن استخدام المعادلات التي تعبر عن القدرة بدلالة العوامل المذكورة سابقاً، أو يمكن استخدام منحنى القدرة وجداول القدرة التي تقدِّر بالفعل حجم العينة. توضح منحنيات وجداول القدرة العلاقة بين القدرة $1-\beta$ والأثر الذي نحاول الكشف عنه. ويمكن لهذه الآثار أن تكون مختلفة في اثنين من المتوسطات، الاختلاف في اثنين من التباينات، ومعامل الارتباط، والاختلاف في متوسطات المعالجة، والاختلافات الأخرى للمجتمعات التي نحاول الكشف عنها من خلال التجربة وتحليل البيانات.

من المهم ملاحظة أن هناك معادلات مختلفة، وبالتالي، منحنيات قدرة مختلفة للآثار المختلفة، التي بدورها يتم اكتشافها من خلال الاختبارات الإحصائية المختلفة. يتم الكشف عن اختلاف في اثنين من المتوسطات (الأثر) بواسطة الاختبار $t$، في حين يتم الكشف عن الاختلاف في اثنين من التباينات (الأثر) باستخدام الاختبار $F$. إن منحنيات القدرة للاختبار $t$ تختلف عن منحنيات القدرة للاختبار $F$.

إن منحنيات القدرة الأكثر استخداماً في تقدير حجم العينة لتجربة طبية حيوية تشمل ما يلي:

١ – اختبار $t$ المزدوج أو غير المزدوج (اختلاف في واحد أو اثنين من متوسطات المجتمعات الإحصائية).

٢ – معامل ارتباط بيرسون (Pearson) (الارتباط بين مجتمعين).

٣ – أنوفا (ANOVA) (الاختلاف في ثلاثة أو أكثر من متوسطات المجتمعات الإحصائية).

ولإجراء اختبار القدرة باستخدام منحنيات القدرة، فإننا بحاجة إلى ما يلي:

١ – حجم الأثر الذي نريد الكشف عنه (أي الاختلاف في اثنين من المتوسطات).

٢ – تقدير بارامترات المجتمع الإحصائي (أي الانحراف المعياري للمجتمع (المجتمعات) استنادا إلى البيانات التجريبية).

٣ – مستوى $\alpha$ (احتمال خطأ النوع I).

٤ – مستوى القدرة = $1-\beta$ (احتمال خطأ النوع II).

يتم اختيار مستويات $\alpha$ و $\beta$ من قبل الباحث قبل تصميم التجربة. كما يتم اختيار حجم الأثر الذي ينبغي الكشف عنه أيضاً من قبل الباحث. يحتاج الباحث مقدماً إلى تحديد ضخامة الأثر لكي يكون مهماً لاستخلاص استنتاجات. وبعبارة أخرى، كم يجب أن يكون اثنان من متوسطات المجتمعات الإحصائية مختلفين بالنسبة إلى إشارة مختلفة شيء

ما في الفسيولوجيا الأساسية، أو آثار الأدوية، أو التغيير في عملية التصنيع. أو كم يجب أن تكون قوة الارتباط لكي تعني شيئاً للباحث، من حيث البيولوجيا الأساسية أو العمليات؟ وبشكل عام، كلما كان الاختلاف الذي يجب الكشف عنه بالاختبار $t$ أصغر أو كلما كانت قيمة الارتباط التي يجب الكشف عنها أكبر، كان حجم العينة أكبر.

قد يدَّعي المرء أن الجزء الأصعب للحصول على المعلومات هو تقدير الانحراف المعياري، أو التباين، للمجتمعات الإحصائية وهذا هو التباين المُفترض للمجتمع الذي يلعب دورًا كبيرًا في تحديد حجم العينة. بشكل عام، كلما كان التباين في المجتمع الإحصائي أكبر، كان حجم العينة المطلوب للحصول على إحصائيات المجتمع الإحصائي أو العملية أكبر، وبالتالي، كان حجم العينة المطلوب لإعطاء قدرة للتحليل الإحصائي أكبر.

يوضح الجدول (٧،١) مثالا عن جدول القدرة للاختبار $t$ غير المزدوج. تم تقدير هذا الجدول عندما $\alpha = 0.05$. ولكن توجد جداول مماثلة لمستويات أخرى لـ $\alpha$. تسمح هذه المنحنيات والجداول بتقدير حجم العينة إذا تم أولاً اختيار الاختلاف الذي تم تسويته في المتوسطين اللذين نريد الكشف عنهما وتحديد القدرة التي يجب عندها إجراء الاختبار $t$. يتم الحصول على الاختلاف الذي يتم تسويته عندما نأخذ الفرق المطلق الذي نود الكشف عنه وتقسيم هذا الفرق على الانحراف المعياري المُقدَّر للمجتمع. مرة أخرى، لدينا تخمين للانحراف المعياري استناداً إلى البيانات التجريبية. عند تسوية الفرق الذي يجب الكشف عنه، فلا داعي للقلق حول وحدات المقياس ويمكن استخدامها من جداول المعايرة.

إن أفضل طريقة لتوضيح استخدام اختبارات القدرة هي العمل من خلال مثال على ذلك. سوف نستخدم الجدول (٧،١) لإيجاد حجم العينة للمثال التالي:

**الهدف:** تحديد ما إذا كان هناك فرق في متوسط الطول بالنسبة للرجال في سن الجامعة والنساء اللواتي يعشن في مدينة المدَّعي؟

**التجربة:** قياس الأطوال في عينات عشوائية من الرجال في سن الجامعة والنساء في مدينة المدَّعي.

**السؤال:** ما عدد العينات الذي نحتاج إليه لرفض فرضية العدم التي تنص على أنه لا يوجد فرق في الطول بين متوسط الرجال في سن الجامعة والنساء في مدينة المدَّعي إذا كان هناك بالفعل فرقٌ حقيقيٌّ في المجتمعين الإحصائيين ؟

**الجدول (١ ,٧). جدول لقدرة الاختبار $t$ غير المزدوج (أحادي الذيل) للطالب ($\alpha = 0.05$).**

| اختلاف في المتوسطات (تم التعبير عنه كنقاط z) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| N | 0.10 | 0.20 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 1.0 | 1.2 |
| 10 | 0.08 | 0.11 | 0.16 | 0.22 | 0.29 | 0.36 | 0.45 | 0.53 | 0.70 | 0.83 |
| 13 | 0.08 | 0.13 | 0.18 | 0.26 | 0.34 | 0.44 | 0.54 | 0.63 | 0.80 | 0.91 |
| 20 | 0.09 | 0.15 | 0.24 | 0.34 | 0.46 | 0.59 | 0.70 | 0.80 | 0.93 | 0.98 |
| 30 | 0.10 | 0.19 | 0.31 | 0.46 | 0.61 | 0.74 | 0.85 | 0.92 | 0.99 | |
| 40 | 0.11 | 0.22 | 0.38 | 0.55 | 0.72 | 0.84 | 0.93 | 0.97 | | |
| 60 | 0.13 | 0.29 | 0.50 | 0.70 | 0.86 | 0.95 | 0.98 | | | |
| 80 | 0.15 | 0.35 | 0.60 | 0.81 | 0.93 | 0.98 | | | | |
| 100 | 0.17 | 0.41 | 0.68 | 0.88 | 0.97 | | | | | |
| 200 | 0.26 | 0.64 | 0.91 | 0.99 | | | | | | |

إذا أردنا الكشف عن الاختلاف الحقيقي في المجتمعات الإحصائية فإننا بحاجة لإجراء اختبار القدرة لتحديد عدد العينات اللازم جمعها من الرجال في سن الجامعة والنساء. إننا نفترض أن المجتمعين الإحصائيين موزَّعان بشكل طبيعي. إن الأثر الذي نحاول اكتشافه هو ما إذا كان هناك اختلاف كبير في متوسطات الطول، وبالتالي، فإننا سوف نستخدم اختبار $t$ غير مزدوج لتحليل البيانات حالما يتم جمعها.

نحن بحاجة قبل إجراء اختبار القدرة إلى اتخاذ قرار بشأن مقدار الأثر الذي نريد الكشف عنه بأنه مهم عندما نقوم بالاختبار $t$. وبالنسبة لهذا المثال سوف نختار فرقًا مقداره 3 بوصة في متوسطات الطول على أنه مهم. ندعي جدلاً أن فرقاً مقداره 3 بوصة في متوسطات المجتمعات الإحصائية هو ذو أهمية بيولوجية.

إن الأثر الذي يجري اختباره $(\mu_1 - \mu_2)$ في هذا المثال هو 3 أنش. لاستخدام المنحنيات الموحَّدة قياسيًّا، نحن بحاجة إلى تسوية هذا الفرق بواسطة تقدير للانحراف المعياري، $\sigma$، للمجتمعات. وهذا هو أساساً النقاط $z$ للفرق بين المتوسطيْن. لاحظ أننا نفترض أن المجتمعين الإحصائيين لهما تباين متساوٍ تقريباً إذا كان علينا استخدام الاختبار $t$. للحصول على تقدير لـ $\sigma$، دعونا نفترض أننا جمعنا بعض البيانات التجريبية وقدرنا الانحراف المعياري للعينة، $s = 5$ بوصة. إن الفرق المعياري الذي تم تسويته ونريد كشفه عندما نطبِّق الاختبار $t$ يكون عندئذ 3/5.

بعد ذلك نختار $\alpha = 0.05$ وقدرة = 0.8. مرة أخرى فإن معدلات الخطأ هذه هي خيارنا.

ومن خلال مقدار الأثر (3/5)، وقيمة القدرة (0.8)، وقيمة $\alpha$ (0.05)، فقد استخدمنا الجداول أو المنحنيات في الجدول (٧،١) لإيجاد عدد العينات المطلوبة من جداول القدرة. على سبيل المثال، نجد أن حجم العينة يجب أن يكون 35 عينة تقريباً.

يمكن إبداء بعض الملاحظات العامة حول تأثير تباين المجتمع، معدلات خطأ النوع I والنوع II، ومقدار التأثير على حجم العينة :

١- بالنسبة لقدرة محددة وتباين معين، فإنه كلما كان الأثر اللازم كشفه أصغر، كان حجم العينة أكبر.

٢- كلما كانت القدرة أكبر، كان حجم العينة أكبر.

٣- كلما كان تباين المجتمع الإحصائي أكبر، كان حجم العينة أكبر.

وإذا استخدمنا نفس المثال السابق والبرنامج Minitab لحساب حجم العينة للاختبار $t$ غير المزدوج عندما نقوم بتغيير العديد من المدخلات لاختبار القدرة، فإننا نجد أحجام العينة التالية (الجدول ٧،٢):

**الجدول رقم (٧،٢). العلاقات بين حجم العينة، والقدرة الإحصائية، والانحراف المعياري، ومقدار الأثر الذي يجب الكشف عنه $(\alpha = 0.05)$.**

| حجم الاختلاف في المتوسطات (بوصة) | الانحراف المعياري (بوصة) | قدرة الاختبار | حجم العينة |
|---|---|---|---|
| 1 | 5 | 0.8 | 394 |
| 2 | 5 | 0.8 | 100 |
| 3 | 5 | 0.8 | 45 |
| 4 | 5 | 0.8 | 26 |
| 3 | 1 | 0.8 | 4 |
| 3 | 2 | 0.8 | 9 |
| 3 | 10 | 0.8 | 176 |
| 3 | 5 | 0.9 | 60 |

وهكذا، يمكننا، قبل إجراء التجربة، تقدير حد أدنى لحجم العينة التي تسمح لنا باستخلاص استنتاجات من البيانات المتوفرة لدينا مع مستوى معين من الثقة والقدرة.

# الفصل الثامن

## مجرد البداية
## Just the Beginning...

قدمت الفصول السبعة السابقة، مقارنة مع معظم الكتب حول الإحصائيات، "مجالاً واعداً" مختصراً حول أساسيات الإحصائيات واستخدامها في التطبيقات الطبية الحيوية. ينبغي للقارئ أن يكون مطلعاً على المفهوم الأساسي لنماذج الاحتمالات، واستخدام مثل هذه النماذج لوصف بيانات العالم الحقيقي، واستخدام الإحصائيات لمقارنة المجتمعات الإحصائية. وينبغي أن يكون القارئ على بينة لتأثير العشوائية والحجب على نتائج التحليل الإحصائي. بالإضافة إلى ذلك، ينبغي أن يكون للقارئ تقديرٌ لأهمية التوزيع الطبيعي في وصف المجتمع الإحصائي، واستخدام الجداول المعيارية، وفكرة أن التحليل الإحصائي يهدف إلى اختبار الفروض بمستوى معين من الثقة. وأخيراً، يجب أن يعرف القارئ أننا نستطيع إيجاد فترات الثقة لأي إحصائية مُقدَّرة.

يغطي هذا الكتاب فقط تلك التحليلات الإحصائية الصالحة لمجتمعات إحصائية أو عمليات يتم نمذجتها بشكل جيد بواسطة التوزيع الطبيعي. بالطبع، هناك العديد من العمليات البيولوجية التي لا يتم نمذجتها بشكل جيد بواسطة التوزيع الطبيعي. وبالنسبة للأنواع الأخرى من التوزيعات، ينبغي للمرء قراءة كتب أكثر تقدماً لمعرفة المزيد عن الإحصائيات غير البارامترية (nonparametric) والإحصائيات عن التوزيعات

أو المجتمعات الإحصائية غير الطبيعية. لا تفترض هذه الاختبارات غير البارامترية توزيعاً أساسياً للبيانات وغالباً ما تكون مفيدة للعينات صغيرة الحجم.

كيف يمكننا تحديد ما إذا كانت البيانات، وبالتالي المجتمع الإحصائي، مُنمذجة بشكل جيد بواسطة التوزيع الطبيعي؟ يمكننا أن نبدأ ببساطة من خلال إيجاد المدرج التكراري للبيانات التي تم أخذ عيناتها فيما يتعلق بالتناظر وتناسب العينات التي تقع داخل واحد، واثنين، وثلاثة، انحرافات معيارية للمتوسط. ويمكننا أيضاً مقارنة متوسط العينة مع وسيط العينة. إن الوسيلة الأكثر رسمية لتحديد مقدار طبيعية عينة هو استخدام اختبار $\chi^2$ [7]. يتضمن اختبار $\chi^2$ مقارنة توزيع العينة الفعلي مع التوزيع الذي سيتم توقعه إذا كانت العينة مأخوذة من توزيع طبيعي. إن الاختلافات بين الترددات المتوقعة لحدوث عينة والترددات الحقيقية لحدوث عينة يتم استخدامها لتقدير إحصائية اختبار $\chi^2$. بعد ذلك يتم مقارنة إحصائية الاختبار هذه مع القيم الحرجة للتوزيع $\chi^2$ لتحديد مستوى الأهمية أو الثقة في رفض فرضية العدم. في هذه الحالة، فإن فرضية العدم هي أن توزيع الترددات (أو نموذج الاحتمال) للتوزيع الأساسي لا يختلف عن التوزيع الطبيعي. ويُشار أيضاً إلى الاختبار $\chi^2$ على أنه اختبار جردة الملاءمة. ونلاحظ أنه يمكن استخدام الاختبار $\chi^2$ أيضاً لمقارنة توزيع عينة مع نماذج احتمال أخرى إلى جانب التوزيع الطبيعي.

لقد قمنا بتغطية أنوفا أحادية وثنائية العامل. ولكن سيواجه المرء أيضاً أنوفا متعدد المتغيرات، التي يوجد فيها أكثر من متغير واحد تابع (أكثر من نتيجة واحدة أو قياس يجري أخذ عيناته). يُشار إلى مثل هذا التحليل باسم مانوفا (MANOVA) ويتم استخدامه في تحليل الارتداد المتعدد الأكثر عمومية الذي قد يكون فيه أكثر من متغير واحد مستقل وأكثر من متغير واحد تابع. تعتبر المتغيرات التابعة مقاييس متعددة (أو مقاييس متكررة) مأخوذة من نفس الوحدات التجريبية التي يجري تعريضها لواحد أو أكثر من العوامل (المتغيرات المستقلة). سعينا في خط الانحدار إلى تحديد مدى إمكانية

توقع سلوك أحد المتغيرات مع الأخذ في الاعتبار سلوك المتغير الثاني. وبعبارة أخرى، كان لدينا متغير واحد مستقل ومتغير واحد تابع. في حالة تحليل الانحدار المتعدد، يمكننا افتراض أكثر من متغير واحد تابع وأكثر من متغير واحد مستقل لتوضيح التباين في المتغير التابع (أو المتغيرات التابعة). على سبيل المثال، قد يكون لدينا متغير واحد تابع، مثل وزن الجسم، الذي نقوم بنمذجته كدالة خطية بثلاثة متغيرات مستقلة: امتصاص السعرات الحرارية، والتمرين، والعمر. عند إجراء تحليل الانحدار المتعدد، فإننا نحاول توقّع مقدار التغير في وزن الجسم بسبب كل واحد من المتغيرات المستقلة الثلاثة. في كثير من الأحيان، فإن متغيراً مستقلاً واحداً ليس كافياً للتنبؤ بنتائج المتغير التابع. ومع ذلك، يجب الأخذ في الاعتبار أنه ينبغي للمرء فقط إضافة أدوات تنبؤ إضافية (المتغيرات المستقلة) تسهم في المتغير التابع بطريقة لا تسهم بها أداة التوقع الأولى. وبعبارة أخرى، يجب على اثنتين من أدوات التنبؤ أو أكثر (المتغيرات المستقلة) توقع النتائج معاً، أو المتغير التابع، بشكل أفضل مما يستطيع أي متغير مستقل أن يتنبأ به بمفرده. لاحظ أن، النموذج الخطي المُعمَّم الكامل يسمح بمتغيرات متعددة مستقلة وتابعة. وفي الجوهر، يبحث نموذج الانحدار المتعدد عن ارتباطات بسيطة بين عدة متغيرات للدخل و الخرج.

على الرغم من أننا ركزنا على خط الانحدار، إلا أن القارئ ينبغي له معرفة أن هناك نماذج للانحدار غير الخطي أيضاً قد تكون قوية جداً في وصف الظواهر البيولوجية. هناك أيضاً أدبيات كاملة عن تحليل الأخطاء التي تنتج عند مقارنة نموذج مُستخدم للتنبؤ بالبيانات، مع البيانات الفعلية المُقاسة [3]. يمكن للمرء في الواقع البحث في الأخطاء المتبقية بين البيانات التي تم التنبؤ بها أو نمذجتها والبيانات المُقاسة فعلاً. إن الشكل المهم أو الاتجاهات في مقدار وترتيب المتبقيات تشير عادة إلى ملاءمة النموذج للبيانات بشكل سيئ. تشير مثل هذه الأشكال إلى أنه ليس كل التغيرات التي يمكن التنبؤ بها في البيانات المقاسة تم أخذها بالاعتبار في النموذج.

تشمل التحليلات المتعددة المتغيرات الأخرى على كل من تحليلات الكتلة، والتمايز، والعامل. ويتم تغطية هذه المواضيع في كتب إحصائية عديدة. تسمح هذه التحليلات بعمل مجتمع إحصائي داخل المجتمع الإحصائي الفرعي وذلك لتوضيح وشرح البيانات التي تبدو معقدة. ومتعددة الأبعاد

لم يتم في هذا الكتاب تغطية المنحنيات المميزة لمشغِّل مستقبل. وهذه هي التحليلات الإحصائية التي يكثر استخدامها في تصميم وتقييم الأجهزة الطبية أو الاختبارات التشخيصية المُستخدمة للكشف عن مرض أو شذوذ. توفر هذه المنحنيات، التي تلخص في كثير من الأحيان مصطلحات مثل الحساسية والنوعية والدقة، وسيلة لتحديد مدى دقة أداة التشخيص، أو الاختبار، أو الخوارزمية في كشف المرض أو الوظيفة الفسيولوجية غير الطبيعية. لقد ناقشنا سابقاً أخطاء النوع I و II، التي يمكن استخدامها أيضاً لتقدير حساسية، ونوعية، ودقة الاختبار التشخيصي. غالباً ما يكون هناك مفاضلة بين الحساسية والنوعية، يمكنها زيادة إحباط المهندس الطبي الحيوي الذي يحاول تطوير اختبارات تشخيصية آمنة ودقيقة وعملية وغير مكلفة. إن المنحنى المميز لمشغِّل مستقبل هو مخطط يرسم حساسية الاختبار (احتمال نتيجة إيجابية صحيحة) مقابل احتمال اختبار إيجابي خاطئ. يختار المُشغِّل عادة العمل على نقطة على المنحنى المميز لمشغِّل مستقبل تكون فيها الحساسية والنوعية عظميين. قد يكون من المفضل في بعض الحالات تخفيض النوعية على حساب الحساسية. على سبيل المثال، عند إدارة اختبار للكشف عن التهاب الحلق بالعقديات، فقد نفضل زيادة الحساسية إلى أقصى حد من أجل عدم تفويت تشخيص بكتيريا الحلق. المفاضلة هي أننا قد نقلل النوعية، ولكن شخصًا ليس لديه التهاب بالعقديات قد يتم تشخيصه خطأً وكأن لديه بكتيريا عقدية. والنتيجة هي أن الشخص ينتهي به المطاف إلى دفع ثمن المضادات الحيوية واستهلاكها دون أن تخدم هذه المضادات أي غرض. وفي حالات سريرية أخرى، قد

تكون المفاضلة بين الحساسية والنوعية أكثر تعقيداً بكثير. على سبيل المثال، نتمنى في كثير من الأحيان استخدام أداة تصوير غير تداخلية (جراحية) للكشف عن آفات الثدي السرطانية. إن الحساسية العالية مطلوبة بحيث لا يضيع الاختبار حدوث آفة سرطانية. ومع ذلك، يجب أن تكون النوعية عالية أيضاً بحيث لا يخضع المريض لعملية جراحية غير ضرورية لإزالة الأنسجة السليمة أو غير السرطانية.

وأخيرًا، هناك مساحة واسعة من التحليل الإحصائي متعلقة بالسلاسل الزمنية، أي البيانات التي يتم جمعها على مر الزمن. إن الزمن متغير مستقل يتم استخدامه للتنبؤ بالمتغير التابع، الذي غالباً ما يكون في التطبيقات الطبية الحيوية عبارة عن مقياس بيولوجي. إن التحليل الإحصائي للسلاسل الزمنية معروض بشكل لطيف في بينديت (Bendat) وبييرسول (Piersol) [7]، ويلعب دوراً هاماً في معظم الأبحاث الطبية الحيوية. وكثيراً ما يُستخدم مثل هذا التحليل لتطوير خوارزميات الكشف المؤتمت (الآلي) لأنظمة مراقبة المرضى، والأجهزة القابلة للزرع، وأنظمة التصوير الطبي.

ويشير هذا الكتاب، ولكن لا يغطي بشكل صريح، إلى استخدام حزم البرامج الإحصائية لتحليل البيانات. هناك عدد من البرامج الإحصائية المتوفرة تجارياً للمهندس الذي يحتاج لإجراء تحليل إحصائي. بعض حزم البرمجيات متاحة مجاناً، في حين أن الأخرى يمكن أن تكون مكلفة جداً. يقدم العديد من حزم البرامج بعض المساعدة التعليمية في استخدام وتفسير التحليل الإحصائي والبياني. بعضها لديه واجهات مستخدم سهلة تسمح للمستخدم بتحميل وتحليل البيانات بسرعة دون الحاجة إلى تعليمات كثيرة. تتطلب الحزم الأخرى قدراً كبيراً من التدريب والممارسة. بعض حزم البرامج الأكثر شعبية للتحليل الإحصائي المُستخدمة في مجال الطب الحيوي تتضمن SPSS، و Minitab، وإكسل (Excel)، و StatView، وماتلاب (Matlab).

وأخيرًا، إن الرسالة الهامة التي يجب أن يحملها القارئ هي أن الإحصائيات ليست كتاب طبخ كما أنها ليست غير منفتحة على التغيير. وحتى أكثر أخصائيي الإحصاء الحيوي خبرة يتناقشون حول أفضل التحليلات التي ينبغي استخدامها للبيانات الطبية الحيوية التي يتم جمعها في ظل ظروف إكلينيكية (سريرية) معقدة. يتطلب التحليل الإحصائي المفيد أن يقوم المستخدم أولاً بوضع فرضية قابلة للاختبار، على أن تُجمع البيانات من عينة تمثيلية، وأن تكون ضوابط التصميم التجريبي للعوامل المربكة عملية قدر الإمكان. حالما يتم جمع البيانات، يحتاج المستخدم إلى أن تكون منطقية قدر الإمكان في تلخيص وتحليل البيانات. وهذا يتطلب أن يكون لدى المستخدم تقدير جيد للفروض الموضوعة حول المجتمعات الإحصائية عند استخدام الاختبار الإحصائي، وكذلك القيود المفروضة على الاختبار الإحصائي في استخلاص استنتاجات محددة. عند استخدام الإحصائيات بشكل صحيح، فإنها تساعدنا بالتأكيد على اتخاذ قرارات جيدة وتنبؤات مفيدة، حتى في سياق عدم اليقين والعوامل العشوائية التي لا نملك السيطرة عليها. يحتاج المهندسون الطبيون الحيويون إلى تبنّي الإحصائيات وتعلمها لكي يكونوا مرتاحين بتطبيق الإحصائيات كما هو الحال بتطبيق الجبر، وحساب التفاضل والتكامل، والمعادلات التفاضلية.

# المراجع
# Bibliography

[1]. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, "Heart rate variability. Standards of measurement, physiologic interpretation and clinical use," Circulation, vol. 93, pp. 1043–1065, 1996.
[2]. Wagner, G.S., Marriott's Practical Electrocardiology, 10th ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2001.
[3]. Hogg, R.V., and Ledolter, J., Engineering Statistics, Macmillan Publishing, New York, 1987.
[4]. Gonick, L., and Smith, W., The Cartoon Guide to Statistics, HarperPerennial, New York, 1993.
[5]. Salkind, N.J., Statistics for People Who (Think They) Hate Statistics, 2nd ed., Sage Publications, Thousand Oaks, CA, 2004.
[6]. Runyon, R.P., Fundamentals of Statistics in the Biological, Medical and Health Sciences, PWS Publishers, Boston, MA, 1985.
[7]. Bendat, J.S., and Piersol, A.G., Random Data. Analysis and Measurement Procedures, 2nd ed., Wiley-Interscience, John Wiley & Sons, New York, 1986.
[8]. Ropella, K.M., Sahakian, A.V., Baerman, J.M., and Swiryn, S., "Effects of procainamide on intra-atrial electrograms during atrial fibrillation: Implications for detection algorithms," Circulation, vol. 77, pp. 1047–1054, 1988.
[9]. Fung, C.A., Descriptive Statistics. Module 1 of a Three-Module Series on Basic Statistics for Abbott Labs, Abbott Labs, Abbott Park, IL, December 11, 1995.
[10]. Olkin, I., Gleser, L.J., and Derman, C., Probability Models and Applications, Macmillan Publishing, New York, 1980.
[11]. Fung, C.A., Comparing Two or More Populations. Module 2 of a Three-Module Series on Basic Statistics for Abbott Labs, Abbott Labs, Abbott Park, IL, December 12, 1995.
[12]. Minitab StatGuide. Minitab Statistical Software, release 13.32, Minitab, State College, PA, 2000.

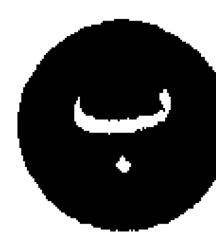
[13]. Ropella, K.M. Baerman, J.M., Sahakian, A.V., and Swiryn, S., "Differentiation of ventricular tachyarrhythmias," Circulation, vol. 82, pp. 2035–2043, 1990.
[14]. Schmit, B.D., Benz, E., and Rymer, W.Z., "Afferent mechanisms of flexor reflexes in spinal cord injury triggered by imposed ankle movements," Experimental Brain Research, vol. 145, pp. 40–49, 2002.

# ثبت المصطلحات

## أولاً: عربي – إنجليزي

### أ

| | |
|---|---|
| إجراء توكي | Tukey's Procedure |
| احتمال | Probability |
| إحصاء | Statistic |
| إحصاء وصفي | Descriptive Statistics |
| اختبار t | t-Test |
| اختبار t المزدوج (المقترن) | Paired t-Test |
| اختبار t غير المزدوج (المقترن) | Unpaired t-Test |
| اختبارات القدرة | Power Tests |
| ارتباط | Correlation |
| ارتباط خطي | Linear Regression |
| انحدار | Regression |
| استنتاج إحصائي | Statistical Inference |



| | |
|---|---|
| بيانات | Data |

## ت

| | |
|---|---|
| تجارب أحادية العامل | One-Factor Experiments |
| تجارب ثنائية العامل | Two-Factor Experiments |
| تجميع البيانات | Collecting Data |
| تحدي | Challenge |
| تحليل | Analysis |
| تحليل إحصائي | Statistical Analysis |
| تحليل ارتباط | Correlation Analysis |
| تحليل القدرة | Power Analysis |
| تصميم تجريبي | Experimental Design |
| تباين | Variance |
| توزيع | Distribution |
| توزيع t | t-Distribution |
| توزيع طبيعي | Normal Distribution |
| توزيع طبيعي معياري | Standard Normal Distribution |

## ث

| | |
|---|---|
| ثقة | Confidence |

## ح

| | |
|---|---|
| حجب | Blocking |
| حجم العينة | Sample Size |

## خ

| | |
|---|---|
| خطي | Linear |

## س

| | |
|---|---|
| سلاسل | Series |
| سلاسل زمنية | Time Series |

## ع

| | |
|---|---|
| عشوائية | randomization |
| عينة | Sample |

## ف

| | |
|---|---|
| فترة | Interval |
| فترة الثقة | Confidence Interval |

## ق

| | |
|---|---|
| قدرة | Power |

## م

| | |
|---|---|
| متوسط | Mean |
| متوسط العينة | Sample Mean |
| متوسطات المجتمع | Population Means |
| مخطط ، رسم | plot |
| المخططات المبعثرة | Scatter plots |
| مخططات الصندوق والمؤشر | Box-and-Whisker plots |
| المدرج التكراري | Histogram |
| معيار، معياري | Standard |
| مقارنة | Comparison |
| مقارنة متعددة | Multiple Comparison |
| مقاييس التغيُّر | Measures of Variability |
| مقاييس النزعة المركزية | Measures of Central Tendency |

## ن

| | |
|---|---|
| نزعة مركزية | Central Tendency |
| نموذج | Model |
| نموذج احتمال | Probability Model |

## هـ

| | |
|---|---|
| هندسة طبية حيوية | Biomedical Engineering |

## ثانياً: إنجليزي– عربي

### A

| | |
|---|---|
| Analysis | تحليل |

### B

| | |
|---|---|
| Biomedical Engineering | هندسة طبية حيوية |
| Blocking | حجب |
| Box-and-Whisker plots | مخططات الصندوق والمؤشر |

### C

| | |
|---|---|
| Central Tendency | نزعة مركزية |
| Challenge | تحدي |
| Collecting Data | تجميع البيانات |
| Comparison | مقارنة |
| Confidence | ثقة |
| Confidence Interval | مدى الثقة |
| Correlation | ارتباط |
| Correlation Analysis | تحليل ارتباط |

### D

| | |
|---|---|
| Data | بيانات |
| Descriptive Statistics | إحصاء وصفي |

| | |
|---|---|
| Distribution | توزيع |

**E**

| | |
|---|---|
| Experimental Design | تصميم تجريبي |

**H**

| | |
|---|---|
| Histogram | المدرج التكرارى |

**I**

| | |
|---|---|
| Interval | مدى |

**L**

| | |
|---|---|
| Linear | خطي |
| Linear Regression | ارتباط خطي |

**M**

| | |
|---|---|
| Mean | متوسط |
| Measures of Central Tendency | مقاييس النزعة المركزية |
| Measures of Variability | مقاييس التغيُّر |
| Model | نموذج |
| Multiple Comparison | مقارنة متعددة |

**N**

| | |
|---|---|
| Normal Distribution | توزيع طبيعي |

## O

| | |
|---|---|
| One-Factor Experiments | تجارب أحادية العامل |

## P

| | |
|---|---|
| Paired t-Test | اختبار t المزدوج (المقترن) |
| Plot | مخطط ، رسم |
| Population Means | متوسطات المجموعات |
| Power | قدرة |
| Power Analysis | تحليل القدرة |
| Power Tests | اختبارات القدرة |
| Probability | احتمال |
| Probability Model | نموذج احتمال |

## R

| | |
|---|---|
| randomization | عشوائية |
| Regression | انحدار |

## S

| | |
|---|---|
| Sample | عينة |
| Sample Mean | متوسط العينة |
| Sample Size | حجم العينة |
| Scatter plots | المخططات المبعثرة |

| | |
|---|---|
| Series | سلاسل |
| Standard | معيار، معياري |
| Standard Normal Distribution | توزيع طبيعي معياري |
| Statistic | إحصاء |
| Statistical Analysis | تحليل إحصائي |
| Statistical Inference | استنتاج إحصائي |

## T

| | |
|---|---|
| t-Distribution | توزيع t |
| Time Series | سلاسل زمنية |
| t-Test | اختبار t |
| Tukey's Procedure | إجراء توكي |
| Two-Factor Experiments | تجارب ثنائية العامل |

## U

| | |
|---|---|
| Unpaired t-Test | اختبار t غير المزدوج (المقترن) |

## V

| | |
|---|---|
| Variance | تباين |

# كشاف الموضوعات

# سيرة المؤلف
# Author Biography

تعمل كريستينا إم روبيلا (Kristina M. Ropella) حالياً أستاذة للهندسة الطبية الحيوية في جامعة ماركيت (Marquette University)، حيث أصبحت عضواً في هيئة التدريس منذ عام ١٩٩٠م. حصلت على درجة دكتوراه فلسفة في الهندسة الطبية الحيوية من جامعة نورث ويسترن (إيفانستون، IL) (Evanston, IL) (Northwestern University). تنحصر اهتماماتها البحثية في معالجة الإشارات الفسيولوجية، والفسيولوجيا الكهربائية للقلب والدماغ، والتصوير بالرنين المغناطيسي الوظيفي، والحوسبة الحيوية (biocomputing). وهي تدرِّس مقررات المرحلة الجامعية والدراسات العليا في معالجة الإشارات الطبية الحيوية، والتحليل الإحصائي للسلاسل الزمنية، والإحصاء الحيوي، والحوسبة الطبية الحيوية، وتصميم الأجهزة الطبية الحيوية. حصلت على جائزة روبرت من جامعة ماركيت (Marquette University Robert) وكلية ماري كيتل (Mary Gettel Faculty) للتميز في التعليم وجائزة كلية الهندسة للمعلم المتميز (College of Engineering Outstanding Teacher award). وهي تدير أيضاً برنامج التصوير الوظيفي، وبرنامج مشترك لدرجة الدكتوراه مُقدَّم من جامعة ماركيت، وكلية الطب في ويسكونسن (Medical College of Wisconsin). حائزة على الزمالة في المعهد الأمريكي للهندسة الطبية والبيولوجية، وهي

عضو بارز في الـ IEEE للهندسة في جمعية الطب والبيولوجيا، وعضو في جمعية الهندسة الطبية الحيوية والجمعية الأمريكية للتعليم الهندسي. وهي عضو في هيئة التحرير لطباعة كتب الـ IEEE، سلسلة الهندسة الطبية الحيوية، وترأست عدداً من جلسات التعليم والجلسات الجامعية ذات الصلة بالصناعة في اجتماعات جمعيات الـ EMBS، والـ BMES، والـ ASEE. تتضمن اهتماماتها الشخصية تربية ثلاثة أطفال، وكرة المضرب، والرقص، والسفر، وواعظة متطوعة للكنيسة.

www.ksu.edu.sa



ISBN 9786035070003